SYMMETRIC MEASURES ON CARTESIAN PRODUCTS[1]

BY

EDWIN HEWITT AND LEONARD J. SAVAGE

1. **Introduction.** This paper has its origin in the theory of probability, but we think it may be of interest to some who are not familiar with probabilistic technique and jargon. Accordingly, we use for the most part the language of measure theory instead of the language of probability. However, an informal probabilitistic statement of the problem will, we hope, pave the way for all readers.

Suppose that, for each value π of a parameter, $\{e_n\}_{n=1}^{\infty}$ is a sequence of random variables that are statistically independent and subject to a common distribution depending on π. If now π itself is a random variable, consider the over-all distribution of the sequence $\{e_n\}_{n=1}^{\infty}$, i.e., the average with respect to π of the conditional distribution of the sequence given π. This over-all distribution will not in general render the e_n's independent, as the conditional distributions given π are assumed to do. Nonetheless, it will obviously, like a distribution with independent e_n's, be invariant under finite permutations of the variables e_n among themselves, or *symmetric*, as we shall say.

Conversely, it is true under very general circumstances that any symmetric distribution on the e_n's can be constructed from a suitable family of independent distributions, parametrized say by π, and a suitable distribution of π.

Jules Haag seems to have been the first author to discuss symmetric sequences of random variables (see [13]). This paper deals only with 2-valued random variables. It hints at, but does not rigorously state or prove, the representation theorem for this case. Somewhat later, symmetric distributions were independently introduced by de Finetti: and the representation theorem for symmetric distributions was proved and exploited by him, especially in connection with the foundations of probability, first in case the e_n's are 2-valued random variables [10; 11]. This case has also been treated by Hinčin, in much the same manner as by de Finetti [18; 19]. De Finetti also proved the theorem for the case in which the e_n's are real random variables [11]. This case has also been treated by Dynkin [9], apparently without knowledge of de Finetti's work in [11]. Dynkin's technique also applies to random variables on all spaces that are, in a certain sense, separable. A few

Presented to the Society, May 2, 1953; received by the editors March 2, 1955.

(1) Research carried out in part at the Statistical Research Center, University of Chicago, under sponsorship of the Statistics Branch, Office of Naval Research, and supported in part by the National Science Foundation.

additional sidelights on the theory and its rôle in the foundations of the theory of probability are mentioned in [24, §3.7]. Another reference, interesting for both its mathematical and historical content, is Chapter III of [12].

The methods of proof thus far presented draw upon devices that appear difficult to extend to any class of random variables e_n not subject to strong separability requirements. We know of no urgent need to achieve such an extension, since even the class of random variables representable as real is very wide from a practical point of view. Still we believe that a more general method of proof may enrich the theory and be of some mathematical interest. The method used in this paper is based on the following idea.

The set of all distributions symmetric on $\{e_n\}_{n=1}^{\infty}$ is obviously convex under linear combination. The representation theorem asserts that every symmetric distribution can be represented as an integral, that is, a sort of generalized linear combination, of independent distributions. Another theorem of the present theory (which appears not to be quite explicit in earlier work, even in the case of real random variables), based on the strong law of large numbers, asserts that this representation is unique. Thus the independent distributions, when the representation theorem is valid, are the extreme points of the convex set of all symmetric distributions. This all suggests that the representation theorem may be reducible to an application of known theorems on the representation of points in convex sets as linear combinations, or the like, of extreme points. The present paper is largely concerned with the implementation of this idea. An announcement of the results has been made in [17].

Many measure-theoretic terms and facts are used throughout this paper; [14] may be consulted for those not explained here. Some knowledge of the theory of extreme points of convex sets is assumed, as set forth, for example, in [4].

To avoid needless repetition, let it be understood that "probability", "algebra", and "isomorphism (homomorphism) between algebras" can mean either *finitely* additive probability, Boolean algebra, and Boolean isomorphism (homomorphism), or *countably* additive probability, σ-algebra, and σ-isomorphism (homomorphism), respectively, with one or the other fixed interpretation throughout. Thus, this paper can be viewed as an abbreviation for 2 papers, either of which can legitimately use deductions made in an earlier part of the other. In specific allusions, we shall call these 2 interpretations the finite and infinite interpretation, respectively, and we shall call their subject matters the finite and infinite situations. A probability will be referred to in terms of the algebra of sets on which it is defined, and, occasionally, as a probability of a certain kind on a certain set.

2. **The problem.** Let $\mathcal{X}$ be an algebra of subsets of a set X; let $\tilde{X}$, with points $a=\{a_n\}_{n=1}^{\infty}$, $b=\{b_n\}_{n=1}^{\infty}$, etc., be the Cartesian product of a countably infinite sequence of replicas of X. For every finite sequence $i_1, \cdots, i_p$ of dis-

tinct positive integers and every finite sequence of sets $E_1, \cdots, E_p$ in $\mathcal{X}$, let $C(E_1^{(i_1)}, \cdots, E_p^{(i_p)})$ denote the set of all $a \in \tilde{X}$ such that $a_{i_r} \in E_i$ $(r=1, \cdots, p)$. Such sets are here called cylinders (rectangles, in [14, p. 154]). For $i_r = r$ $(r=1, \cdots, p)$, we shall write $C(E_1^{(i_1)}, \cdots, E_p^{(i_p)})$ as $C(E_1, \cdots, E_p)$; any cylinder can clearly be expressed in this simpler form. We now define $\tilde{\mathcal{X}}$ as the smallest algebra of subsets of $\tilde{X}$ containing all cylinders.

Let $\boldsymbol{P}$ be the set of all probabilities π on $\mathcal{X}$, and let $\tilde{\boldsymbol{P}}$ be the set of corresponding product measures $\tilde{\pi}$ on $\tilde{\mathcal{X}}$ (see [14, p. 157, Theorem B]). The correspondence $\pi \to \tilde{\pi}$ is clearly 1-to-1 onto.

The set $\tilde{\mathcal{S}}$ of symmetric probabilities on $\tilde{\mathcal{X}}$ is defined as follows. Let T be an arbitrary 1-to-1 mapping of the positive integers onto themselves leaving all but a finite number of integers fixed. For $A \subset \tilde{X}$, let TA be the set of all $\{a_n\}_{n=1}^{\infty} \in \tilde{X}$ such that $\{a_{T(1)}, a_{T(2)}, \cdots, a_{T(n)}, \cdots\} \in A$ (though there would be good reason to call this $T^{-1}A$). A probability σ on $\tilde{\mathcal{X}}$ is in $\tilde{\mathcal{S}}$, and is said to be *symmetric*, if and only if $\sigma(TA) = \sigma A$ for all $A \in \tilde{\mathcal{X}}$ and for all T. It is obvious that $\tilde{\boldsymbol{P}} \subset \tilde{\mathcal{S}}$.

The property of symmetry of a probability σ on $\tilde{\mathcal{X}}$ can be defined in a different way. Let n be an arbitrary positive integer and let $E_1, \cdots, E_n$ be any sets in $\mathcal{X}$. It is plain that σ is symmetric if and only if

$$\sigma C(E_1^{(i_1)}, \cdots, E_n^{(i_n)}) = \sigma C(E_1^{(j_1)}, \cdots, E_n^{(j_n)}) \tag{2.1}$$

for all sequences $i_1, \cdots, i_n$ and $j_1, \cdots, j_n$ of positive integers (i's all distinct and j's all distinct).

With a view toward integrating over the set $\boldsymbol{P}$, we define a σ-algebra $\mathcal{P}^*$ of subsets of $\boldsymbol{P}$ as follows. For all real numbers λ and all $E \in \mathcal{X}$, the set

$$\{\pi;\ \pi \in \boldsymbol{P},\ \pi(E) \leqq \lambda\} = \boldsymbol{N}(E; \lambda) \tag{2.2}$$

is an element of $\mathcal{P}^*$, and $\mathcal{P}^*$ is the smallest σ-algebra of subsets of $\boldsymbol{P}$ containing all sets of the form (2.2).

It may be the case that a given $\sigma \in \tilde{\mathcal{S}}$ can be represented as an average of π's in $\boldsymbol{P}$ in the sense that there exists a countably additive probability μ on $\mathcal{P}^*$ such that

$$\sigma A = \int_{\boldsymbol{P}} (\tilde{\pi} A)\, d\mu(\pi) \tag{2.3}$$

for all $A \in \tilde{\mathcal{X}}$. Such probabilities σ are said to be *presentable*.

The 1-to-1 correspondence $\pi \leftrightarrow \tilde{\pi}$ between $\boldsymbol{P}$ and $\tilde{\boldsymbol{P}}$ induces an isomorphism between the σ-algebra $\mathcal{P}^*$ and a σ-algebra, say $\tilde{\mathcal{P}}^*$, of subsets of $\tilde{\boldsymbol{P}}$. Similarly, a probability μ on $\mathcal{P}^*$ has an image $\tilde{\mu}$ which is a probability on $\tilde{\mathcal{P}}^*$. Hence (2.3) is equivalent to

$$\sigma A = \int_{\tilde{\boldsymbol{P}}} (\tilde{\pi} A)\, d\tilde{\mu}(\tilde{\pi}). \tag{2.4}$$

The formula (2.4) indicates that a presentable probability σ is, in a certain sense, a *mixture* of elements of $\tilde{P}$.

If the algebra $\mathcal{X}$ has the property that every probability σ in $\tilde{\mathbf{S}}$ is presentable, we shall say that $\mathcal{X}$ is presentable. The object of the present paper is to find conditions under which $\mathcal{X}$ is presentable. In the infinite situation, we find that a fairly extensive class of σ-algebras is presentable, but the question of the existence of unpresentable $\mathcal{X}$ is, unfortunately, left unsolved. In the finite situation, we prove that *every* Boolean algebra is presentable.

3. **Orientation.** To clarify the study of presentability, numerous assertions are made in this section proofs of which are explicitly either deferred to later sections or omitted as trivial.

For $\sigma \in \tilde{\mathbf{S}}$ to be presentable, it is necessary, in view of (2.3), that $\tilde{\pi}A$ be $\mathcal{P}^*$-measurable as a function of π for every $A \in \mathcal{X}$; this condition is always satisfied (Theorem 4.1).

For every countably additive probability measure μ on $\mathcal{P}^*$, (2.3) does define an element of $\tilde{\mathbf{S}}$ (Theorem 4.3). The algebra $\mathcal{X}$ is presentable, then, if and only if the subset of $\tilde{\mathbf{S}}$ consisting of all measures (2.3) is actually all of $\tilde{\mathbf{S}}$.

If $\sigma \in \tilde{\mathbf{S}}$ is presentable, there is a unique measure μ on $\mathcal{P}^*$ for which (2.2) holds, and it is given by an explicit algorithm (Theorems 9.4 and 9.2).

In view of the trivial fact that presentability of $\mathcal{X}$ is invariant under isomorphisms of $\mathcal{X}$, it would seem a natural possibility to develop all of the present theory for an abstract Boolean algebra $\mathcal{X}$ rather than for an algebra of sets. By Stone's representation theorem for Boolean algebras [26], every abstract Boolean algebra is isomorphic to an algebra of sets, and thus, in the finite situation, nothing is gained by considering abstract algebras *per se*. In the infinite situation, the matter is more complicated, because there exist familiar abstract σ-algebras that are σ-isomorphic to no algebra of sets [21], and for other reasons. We have done some work on the subject, but it is rather foreign to the spirit of the present paper, so we hope to publish it separately. Briefly, there is a strong tendency for an abstract σ-algebra that can be called presentable to imitate a σ-algebra of sets. The remark following the proof of Theorem 11.7 hints at this phenomenon.

The alert reader will have noted that we require the measure μ that figures in (2.3) to be countably additive. By admitting finitely additive rather than only countably additive μ's in (2.3), we would get a class of symmetric measures that might be called *pseudo-presentable*—in the infinite situation, not typically one for every μ. This concept has no interest in the finite situation, since every symmetric probability σ is presentable (Theorem 8.1) and satisfies (2.3) for only one measure μ, which is already countably additive. If there exists a countably additive symmetric probability σ that is unpresentable, it would still be of interest to know that (2.3) is satisfied for this σ with a *finitely* additive μ. Note that such a μ would still have to be defined not merely on the Boolean algebra generated by (2.2) but on the entire

σ-algebra $\mathcal{P}^*$, as is easily seen from Theorem 4.2 and the fact that $\tilde{\pi}A$ must be measurable as a function of π.

The problem under study can be generalized by admitting uncountably infinite Cartesian products of the basic set X and also by admitting finite Cartesian products of X. The former generalization is of no interest at all (as Dynkin has pointed out [9]), since every symmetric probability on an arbitrary infinite Cartesian product is obviously determined by its values for sets restricted only on any one countably infinite set of co-ordinates. On the other hand, finite products are essentially different from infinite products. §12 is devoted to a study of these products, and it is shown that symmetric probabilities on finite products all have a property akin to presentability.

We now make a few remarks about cylinders. Since the intersection of 2 cylinders is plainly a cylinder and since the equality

$$(3.1) \quad C(E_1, \cdots, E_n)' = C(E_1, \cdots, E_{n-1})' \cup C(E_1, \cdots, E_{n-1}, E_n')$$

holds for all $n \geqq 1$ and $E_1, \cdots, E_n$, we see at once that the family $\mathcal{C}$ of all cylinders forms a semiring [14, p. 22, exercise 4.6]. This semiring includes $\tilde{X}$ and may hence be called a semialgebra.

The class of cylinders is not the only semialgebra with which we have to deal, and some facts about semialgebras will be useful. These facts are summarized in Theorem 3.1 below, the various parts of which are well known or easy to prove (cf. [14, p. 26, ex. 3, p. 37, ex. 5, and p. 57, ex. 1]). The concepts of probability, homomorphism, and isomorphism are applied to semialgebras in the obvious senses. It should, however, be emphasized that, for a measure or homomorphism on a semialgebra, countable additivity is not implied by good behavior on descending sequences, so that in applications countable additivity must be verified explicitly (cf. [14, p. 214, ex. 3]).

THEOREM 3.1. *Let $\mathcal{C}$ be a semialgebra of subsets of a set X, and let $\mathfrak{X}$ be the smallest algebra of subsets of X containing $\mathcal{C}$. Every element of the Boolean algebra generated by $\mathcal{C}$ can be written (not necessarily uniquely) as the union of a finite number of pairwise disjoint elements of $\mathcal{C}$. Let γ (f) be a probability (homomorphism) on $\mathcal{C}$. Then γ (f) admits a unique extension over $\mathfrak{X}$ that is also a probability (homomorphism). If f on $\mathcal{C}$ is an isomorphism, then its extension is also an isomorphism.*

We leave the proof of this theorem to the reader, who may turn to the exercises in [14] referred to above.

A measure $\sigma \in \tilde{\mathbf{S}}$ is, by definition, invariant under finite permutations of the co-ordinates. We shall now show that every $\sigma \in \tilde{\mathbf{S}}$ is invariant under *all* permutations of the co-ordinates and even under a somewhat wider class of transformations than that.

THEOREM 3.2. *Let T be a* 1-to-1 *transformation of any infinite subset of the positive integers onto all of the positive integers, and define TA as in* §2. *Every T*

is an isomorphism of $\tilde{\mathcal{X}}$ into (not necessarily onto) itself, and $\sigma TA = \sigma A$ for all T, all $\sigma \in \tilde{\mathbf{S}}$, and all $A \in \tilde{\mathcal{X}}$.

Proof. For all $A \subset \bar{X}$, it is obvious that $(TA)' = T(A')$, and for all $\{A_i\}_{i \in I}$ $(A_i \subset \bar{X})$, it is obvious that $\bigcup_i TA_i = T(\bigcup_i A_i)$. Hence T is a σ-homomorphism of the family of all subsets of $\bar{X}$ into itself. But this homomorphism T clearly carries nonvoid sets into nonvoid sets and is, therefore, an isomorphism. It is also clear that TC is a cylinder if C is a cylinder, and that $T\tilde{\mathcal{X}} \subset \tilde{\mathcal{X}}$.

Now consider the probability σ confined to the semialgebra of cylinders. Here, by (2.1) we have $\sigma TC = \sigma C$. The set-function σT, defined at A as $\sigma(TA)$, is an extension of σ and σT to the algebra generated by cylinders. By Theorem 3.1, we have $\sigma = \sigma T$ on this algebra; that is, $\sigma A = \sigma TA$ for all $A \in \tilde{\mathcal{X}}$.

It may be well to mention that sets invariant under finite permutations are not ordinarily invariant under all permutations. For example, sets determined by limiting conditions obviously belong to the first class but not to the second. Indeed, sets invariant under all permutations are easily seen to be of a simple and uninteresting structure.

4. **Measurable sets in $\tilde{P}$.** The σ-algebra $\tilde{\mathcal{P}}^*$ of subsets of $\tilde{P}$ was defined in §2 as an image of the σ-algebra $\mathcal{P}^*$. A 2nd algebra of subsets of $\tilde{P}$ can be defined by paraphrasing the definition of $\mathcal{P}^*$ in terms of $\tilde{\pi}A$, $A \in \tilde{\mathcal{X}}$, rather than in the original terms of πE, $E \in \mathcal{X}$. These 2 algebras coincide, as the 2nd of the following theorems shows.

THEOREM 4.1. *Let A be any set in $\tilde{\mathcal{X}}$. The function f on $\tilde{P}$ defined by $f(\tilde{\pi}) = \tilde{\pi}A$ is $\tilde{\mathcal{P}}^*$-measurable.*

Proof. Let $\tilde{\mathcal{M}}$ denote the family of all $A \in \tilde{\mathcal{X}}$ for which the conclusion of the present theorem holds. If $E \in \mathcal{X}$, then $C(E) \in \tilde{\mathcal{M}}$, since $\tilde{\pi}C(E) = \pi E$, and the function πE is $\mathcal{P}^*$-measurable. Since $\pi C(E_1, E_2, \cdots, E_n) = \pi C(E_1) \cdot \pi C(E_2) \cdot \cdots \cdot \pi C(E_n)$, it follows that $C(E_1, E_2, \cdots, E_n) \in \tilde{\mathcal{M}}$. Since $\tilde{\pi}(A \cup B) = \tilde{\pi}A + \tilde{\pi}B - \tilde{\pi}(A \cap B)$, it follows that $\tilde{\mathcal{M}}$ is closed under the formation of finite pairwise disjoint unions. Therefore $\tilde{\mathcal{M}}$ contains all finite pairwise disjoint unions of cylinders; this family of sets is a Boolean algebra (Theorem 3.1) that generates $\tilde{\mathcal{X}}$. In the finite situation, this implies that $\tilde{\mathcal{M}} = \tilde{\mathcal{X}}$, as was to be proved. To complete the proof in the infinite situation, let $\{A_k\}_{k=1}^{\infty}$ be any monotone infinite sequence of sets in $\tilde{\mathcal{M}}$. Then

$$\lim_{k \to \infty} A_k \in \tilde{\mathcal{M}},$$

since $\tilde{\pi}(\lim_{k \to \infty} A_k) = \lim_{k \to \infty} \tilde{\pi}(A_k)$, and limits of measurable functions are measurable, when measurability is defined with respect to a σ-algebra of sets. It now follows that $\tilde{\mathcal{M}} = \tilde{\mathcal{X}}$, since $\tilde{\mathcal{M}}$ is closed under the formation of monotone limits and contains a Boolean algebra that generates $\tilde{\mathcal{X}}$.

THEOREM 4.2. *Let $\tilde{\mathcal{P}}^{**}$ be the smallest σ-algebra of subsets of $\tilde{P}$ containing all sets of the form*

$$\{\tilde{\pi} \mid \tilde{\pi} \in \tilde{P}, \tilde{\pi}A \leqq \lambda\}, \tag{4.1}$$

for $A \in \tilde{\mathcal{X}}$ and real λ. Then $\tilde{\mathcal{P}}^{**} = \tilde{\mathcal{P}}^{*}$.

Proof. Theorem 4.1 implies immediately that $\tilde{\mathcal{P}}^{**} \subset \tilde{\mathcal{P}}^{*}$. Since $\tilde{\pi}C(E) = \pi E$ for all $E \in \mathcal{X}$, the reverse inclusion also holds.

The following assertion is an easy consequence of Theorem 4.1 (and, in the infinite situation, Lebesgue's theorem on monotone convergence).

THEOREM 4.3. *Let μ be any countably additive probability defined on $\mathcal{P}^{*}$. Then the set-function σ defined on $\tilde{\mathcal{X}}$ by the equality* (2.3) *is a symmetric probability.*

5. **The extreme points of $\tilde{S}$.** In the present section, we characterize the extreme points of the convex set $\tilde{S}$ as being exactly the elements of $\tilde{P}$. The purpose of this characterization is to prepare for theorems showing that many algebras $\mathcal{X}$ (in the finite situation, *all* $\mathcal{X}$) are presentable. It will be obvious that every theorem of the present section is itself a consequence of presentability of $\mathcal{X}$, and is therefore of temporary interest only, at least for presentable $\mathcal{X}$. The proofs may appear somewhat contrived, but their motivation will, we hope, be clear to anyone used to working with probability, especially if he keeps in mind the model of a presentable algebra $\mathcal{X}$.

The class $\tilde{S}$ of symmetric probabilities on $\tilde{\mathcal{X}}$ is obviously a convex set in the linear space of all finite measures on $\tilde{\mathcal{X}}$ in the familiar sense that the set-function σ defined by the relation

$$\sigma = \alpha\sigma' + (1 - \alpha)\sigma'' \tag{5.1}$$

is an element of $\tilde{S}$ provided that σ', σ'' are elements of $\tilde{S}$ and $0 \leqq \alpha \leqq 1$. If $0 < \alpha < 1$, then any 3 measures σ, σ', and σ'' satisfying (5.1) are obviously all identical or else all distinct. As usual, we say that $\sigma \in \tilde{S}$ is an extreme point of $\tilde{S}$ if σ cannot be represented in the form (5.1) with $0 < \alpha < 1$ and $\sigma' \neq \sigma''$.

THEOREM 5.1. *Let n be a positive integer, let $E_1, E_2, \cdots, E_n$ be elements of $\mathcal{X}$, and let σ be an element of $\tilde{S}$. Then*

$$\sigma C(E_1, E_2, \cdots, E_n, E_1, E_2, \cdots, E_n) \geqq [\sigma C(E_1, E_2, \cdots, E_n)]^2. \tag{5.2}$$

Proof. Let the cylinders appearing on the left and right sides of (5.2) be denoted by A and B, respectively, and let χ_r ($r = 1, 2, 3, \cdots$) be the characteristic function of the cylinder

$$\{a \mid a \in \tilde{X}, a_{i+(r-1)n} \in E_i, i = 1, 2, \cdots, n\}. \tag{5.3}$$

Then, for every positive integer m, it is easy to see that

$$\int_{\tilde{X}} \left[\sum_{r=1}^{m} \chi_r(a) \right] d\sigma(a) = m\sigma B, \tag{5.4}$$

since σ is symmetric. Furthermore, a similar direct calculation shows that

$$\begin{aligned} &\int_{\tilde{X}} \left[\sum_{r=1}^{m} \chi_r(a) \right]^2 d\sigma(a) \\ &\quad = \sum_{r=1}^{m} \sum_{s=1}^{m} \int_{\tilde{X}} [\chi_r(a)\chi_s(a)]\, d\sigma(a) \\ &\quad = m \int_{\tilde{X}} \chi_1(a)\, d\sigma(a) + m(m-1) \int_{\tilde{X}} [\chi_1(a)\chi_2(a)]\, d\sigma(a) \\ &\quad = m\sigma B + m(m-1)\sigma A. \end{aligned} \tag{5.5}$$

Applying the Cauchy-Schwarz inequality, $(\int fg)^2 \leqq (\int f^2)(\int g^2)$, with $f = \sum_{r=1}^{m} \chi_r$ and g identically 1, and noting that σ is a probability (i.e., applying the probabilistic fact that no variance is negative), we have

$$\left\{ \int_{\tilde{X}} \left[\sum_{r=1}^{m} \chi_r(a) \right] d\sigma(a) \right\}^2 \leqq \int_{\tilde{X}} \left[\sum_{r=1}^{m} \chi_r(a) \right]^2 d\sigma(a). \tag{5.6}$$

Combining (5.4), (5.5), and (5.6), we find

$$\sigma A \geqq (\sigma B)^2 - \frac{1}{m}(\sigma B - \sigma A) \geqq (\sigma B)^2 - \frac{1}{m}. \tag{5.7}$$

Since (5.7) holds for all positive integers m, the present theorem is proved.

THEOREM 5.2. *Let σ be an element of $\tilde{\mathbf{S}}$ such that equality obtains in* (5.2) *for all positive integers n and all sets $E_1, E_2, \cdots, E_n \in \mathfrak{X}$. Then σ is an extreme point of $\tilde{\mathbf{S}}$.*

Proof. If $\sigma \in \tilde{\mathbf{S}}$ and σ is not an extreme point, there exist $\sigma', \sigma'' \in \tilde{\mathbf{S}}$ and α, $0 < \alpha < 1$, such that $\sigma' \neq \sigma''$ and $\sigma = \alpha\sigma' + (1-\alpha)\sigma''$. Since all measures on $\tilde{\mathfrak{X}}$ are determined by their values on cylinders (Theorem 3.1), there exists a cylinder $B = C(E_1, E_2, \cdots, E_n)$ such that $\sigma' B \neq \sigma'' B$. Let

$$A = C(E_1, E_2, \cdots, E_n, E_1, E_2, \cdots, E_n).$$

Then we have

$$\begin{aligned} \sigma A &= \alpha\sigma' A + (1-\alpha)\sigma'' A \\ &\geqq \alpha(\sigma' B)^2 + (1-\alpha)(\sigma'' B)^2, \end{aligned} \tag{5.8}$$

in view of Theorem 5.1. Applying the Cauchy-Schwarz inequality again, we have

$$[\alpha\sigma' B + (1-\alpha)\sigma'' B]^2 < \alpha(\sigma' B)^2 + (1-\alpha)(\sigma'' B)^2, \tag{5.9}$$

since the conditions for equality obviously fail here. Combining (5.8) and (5.9), we obtain

$$\sigma A > [\alpha\sigma' B + (1 - \alpha)\sigma'' B]^2 = (\sigma B)^2, \tag{5.10}$$

so that strict inequality holds in (5.2). This proves the present theorem.

To prove the next theorem, and to carry out some later computations, the following notation is useful. For a probability π on an algebra Υ of subsets of a set Y, let $E \in \Upsilon$ be such that $\pi E \neq 0$. Then the set function $\pi \mid E$ defined on $F \in \Upsilon$ by the relation

$$\pi \mid E(F) = \pi(E \cap F)/\pi E \tag{5.11}$$

is a probability on Υ. It is usually called the conditional probability given E.

THEOREM 5.3. *$\tilde{P}$ is the set of extreme points of $\tilde{S}$.*

Proof. Let $\tilde{\pi}$ be any element of $\tilde{P}$. Then we obviously have equality in (5.2) for the probability $\tilde{\pi}$, and it follows from Theorem 5.2 that $\tilde{\pi}$ is an extreme point of $\tilde{S}$.

To prove the reverse inclusion, suppose that σ is an element of $\tilde{S}$ that is not in $\tilde{P}$. Since σ is symmetric, it cannot then be a product measure at all. Accordingly, there must exist sets $E, F_1, F_2, \cdots, F_n \in \mathcal{X}$ such that the non-equality

$$\sigma C(E, F_1, F_2, \cdots, F_n) \neq \sigma C(E)\sigma C(F_1, F_2, \cdots, F_n) \tag{5.12}$$

holds. Now, for all $A \subset \tilde{X}$, let UA be the set

$$\{a \mid a \in \tilde{X}, \{a_2, a_3, \cdots\} \in A\}. \tag{5.13}$$

Crudely speaking, UA is the set obtained from A by pushing A one coordinate to the right. Theorem 3.2 implies that the transformation $A \to UA$ is an isomorphism of $\tilde{\mathcal{X}}$ into (not, in general, onto) $\tilde{\mathcal{X}}$, and that $\sigma UA = \sigma A$ for all $A \in \tilde{\mathcal{X}}$.

The condition (5.12) can be rephrased in terms of U as follows. There are a set $B = C(F_1, F_2, \cdots, F_n) \in \tilde{\mathcal{X}}$ and a set $E \in \mathcal{X}$ such that

$$\sigma[C(E) \cap UB] \neq \sigma C(E) \cdot \sigma B. \tag{5.14}$$

In view of (5.14), it is impossible that $\sigma C(E)$ or $\sigma C(E')$ vanish. Therefore we can define set-functions σ' and σ'' for $A \in \tilde{\mathcal{X}}$ as follows:

$$\sigma' A = \sigma \mid C(E)(UA), \qquad \sigma'' A = \sigma \mid C(E')(UA). \tag{5.15}$$

It is easy to see that σ' and σ'' are distinct elements of $\tilde{S}$, in view of (5.14), and that

$$\sigma = [\sigma C(E)]\sigma' + [1 - \sigma C(E)]\sigma''. \tag{5.16}$$

Therefore, if $\sigma \in \tilde{S}$ and σ non$\in \tilde{P}$, then σ is not an extreme point of $\tilde{S}$. This completes the present proof.

6. **Averages of extreme points.** In the present section, we show that all elements of certain convex sets can be represented as averages (that is, countably additive probability integrals) over the extreme points of the convex sets under discussion. The first theorem of the present section has also been proved by Tomita [27], and a generalization of it by Bourbaki [5, p. 87, Proposition 7]. The proof given here seems simpler than Bourbaki's, and is more accessible than Tomita's.

Let L be a normed linear space over the real numbers, with elements f, f', $\cdots$. Let L^*, with elements g, g', $\cdots$ be the conjugate space of L, i.e., the space of all linear functionals on L that are continuous in the topology defined by the norm of L. It is well known and obvious that L^* is a Banach space.

Suppose next that W is a measure space under the countably additive probability μ and that g is a function with domain W and range contained in L^*. We shall say that g is *μ-integrable* if the real-valued functions defined on W by the products (f, g) are μ-integrable for all $f \in L$ and if

$$\int_W (f, g(w))d\mu(w), \tag{6.1}$$

considered as a linear functional on L, is bounded on the set $\|f\| \leqq 1$. The linear functional defined by (6.1) is in this case an element of L^*; we denote it by the symbol

$$\int_W g(w)d\mu(w). \tag{6.2}$$

It is defined, of course, by the condition

$$\left(f, \int_W g(w)d\mu(w)\right) = \int_W (f, g(w))d\mu(w). \tag{6.3}$$

We shall not dilate on the properties of the integral $\int_W g(w)dw$; they can be easily developed following Pettis [23]. We shall apply this integral here only in the very special case where W is a bounded subset M of L^* itself, with measurable sets in W defined by the weak Baire sets $\mathfrak{M}$ of M, and with the (obviously integrable) function g defined as the identity mapping of W into L^*. In this context, the required properties of $\int_W g(w)d\mu(w)$ are simple to verify.

Suppose, then, that M is a bounded subset of L^*. Let $I(M)$ be defined as

$$\left\{g' \mid g' \in L^*, g' = \int_M g d\mu(g),\ \mu \text{ a countably additive probability on } \mathfrak{M}\right\}. \tag{6.4}$$

For every $N \subset L^*$, let $K(N)$ be the smallest weakly closed convex subset of L^* containing N.

THEOREM 6.1. *Let M be a bounded and weakly closed subset of L^*. Then $I(M)=K(M)$.*

Proof. It is obvious that $I(M)$ is convex. It is also clear, by using 1-point probabilities on $\mathfrak{M}$ (that is, measures equal to 1 for every measurable set containing a given point) that $I(M)\supset M$. To show that $I(M)$ is weakly closed, consider the space $\mathfrak{C}(M)$ of all weakly continuous real-valued functions on M. The set of all probabilities μ on $\mathfrak{M}$ constitutes, as is well known, a compact subset of the conjugate space $\mathfrak{C}^*(M)$ in its weak topology. The mapping $\mu\to\int_M g\,d\mu(g)$ is a continuous mapping of this subset of $\mathfrak{C}^*(M)$ onto $I(M)$, where $I(M)$ is given its relative topology as a subset of L^* in its weak topology; for neighborhoods in the weak topology of L^* are based on the special continuous functions on L^* of the form (f, g) for $f\in L$, and the inverse images of such neighborhoods are easily seen to be neighborhoods in the space of μ's regarded as a subset of $\mathfrak{C}^*(M)$. Continuous images of compact spaces being compact, it follows that $I(M)$ is compact in the weak topology of L^*. Since a compact subspace of a Hausdorff space is closed, $I(M)$ is weakly closed.

The foregoing discussion shows that $I(M)\supset K(M)$. In order to show that $I(M)\subset K(M)$, consider an element g_0 of L^* that is not in $K(M)$. Since $K(M)$ is regularly convex (see for example [20]), there exists $f\in L$ such that

$$(f, g_0) > \max_{g\in K(M)} (f, g). \tag{6.5}$$

However, if $g_0\in I(M)$, we have by definition that

$$\begin{aligned} (f, g_0) &= \int_M (f, g)\,d\mu(g) \\ &\leqq \max_{g\in M} (f, g) \\ &\leqq \max_{g\in K(M)} (f, g). \end{aligned} \tag{6.6}$$

It follows from this contradiction that g_0 non$\in I(M)$. This completes the proof.

THEOREM 6.2. *Let N be a weakly compact, convex subset of L^*, let P be the set of extreme points of N, and let P^- be the weak closure of P. Then $N=I(P^-)$.*

Proof. According to the Kreĭn-Mil'man theorem [20], we have $N=K(P)$. Since $K(P)\subset K(P^-)\subset K(N)=N$, we also have $N=K(P^-)$. Finally, Theorem 6.1 implies that $K(P^-)=I(P^-)$.

THEOREM 6.3. *Let N and P be as in Theorem* 6.2, *and suppose further that P is weakly closed. Then $N=I(P)$.*

Proof. This is a trivial corollary of Theorem 6.2.

The condition $P=P^-$ is automatically satisfied in no L^* of dimension >2. See for example [4, p. 87, ex. 8]. In finite dimensional spaces, however, it is the case that $I(P)=N$ even if P fails to be closed. For, if dim $N=n$, where n is a non-negative integer, then every point in N is a convex combination of at most $n+1$ extreme points of N. (See for example [3, p. 9]).

THEOREM 6.4. *Let N be a bounded and weakly closed subset of L^*. Then the Baire sets of N in its weak topology are the σ-algebra $\mathcal{J}$ generated by sets of the form*

$$\{g \mid g \in N, (f, g) \leqq \lambda\} \tag{6.7}$$

for arbitrary $f \in L$ and real λ.

Proof. Since (f, g) is a weakly continuous function of g, it follows that $\mathcal{J}$ is contained in the family of Baire sets. To prove the reverse inclusion, note that every function (f, g) on N is $\mathcal{J}$-measurable. Therefore polynomials in such functions are $\mathcal{J}$-measurable, and by the Stone-Weierstrass theorem [25, p. 466, Theorem 82], all continuous functions on N are $\mathcal{J}$-measurable. Hence $\mathcal{J}$ contains the family of Baire sets.

7. **Sufficient conditions for presentability in the infinite situation.** The present section refers only to the infinite situation.

THEOREM 7.1. *Let X be a compact Hausdorff space, and let $\mathfrak{X}$ be the σ-algebra of all Baire sets in X. Then $\tilde{\mathfrak{X}}$ is the family $\tilde{\mathcal{B}}$ of all Baire sets under the Cartesian product topology of $\tilde{X}$.*

Proof. Given a continuous real-valued function f on X and a positive integer k, the function g on $\tilde{X}$ defined by

$$g(a) = f(a_k) \quad \text{for all} \quad a \in \tilde{X} \tag{7.1}$$

is evidently continuous on $\tilde{X}$. Therefore, if E is a compact G_δ in X and k is any positive integer, the cylinder $C(E^{(k)})$ is an element of $\tilde{\mathcal{B}}$, since $C(E^{(k)}) = \{a \mid g(a)=0\}$ for some function g of the form (7.1). If E is any set in $\mathfrak{X}$, it follows from this that $C(E^{(k)})$ is in $\tilde{\mathcal{B}}$. Hence all cylinders are in $\tilde{\mathcal{B}}$, and therefore, since $\tilde{\mathcal{B}}$ is a σ-algebra, we have $\tilde{\mathfrak{X}} \subset \tilde{\mathcal{B}}$. To prove the reverse inclusion, note first that polynomials in functions of the form (7.1) are uniformly dense in the space of all continuous real-valued functions on $\tilde{X}$. Since all functions (7.1) are obviously $\tilde{\mathfrak{X}}$-measurable, it follows that all continuous functions on $\tilde{X}$ are $\tilde{\mathfrak{X}}$-measurable, and thus that $\tilde{\mathcal{B}} \subset \tilde{\mathfrak{X}}$.

Now, retaining the hypotheses of Theorem 7.1, we consider the Banach space $\mathfrak{C}(\tilde{X})$ of all continuous real-valued functions on $\tilde{X}$, with the usual algebraic operations and with the norm

$$\|f\| = \max_{a \in \tilde{X}} |f(a)|. \tag{7.2}$$

According to the representation theorem of F. Riesz (see for example [14,

p. 247, Theorem D]), the conjugate space $\mathfrak{C}^*(\tilde{X})$ admits a concrete representation as the space of all countably additive, real, finite-valued Baire measures on $\tilde{\mathcal{B}}$, that is, such measures defined on $\tilde{\mathcal{X}}$ (Theorem 7.1). For every such measure ϕ, the corresponding linear functional is defined for all $f \in \mathfrak{C}(\tilde{X})$ by the relation

$$(f, \phi) = \int_{\tilde{X}} f(a) d\phi(a). \tag{7.3}$$

It is an elementary exercise, which we omit, to prove that the norm $\|\phi\|$ of ϕ considered as an element of $\mathfrak{C}^*(\tilde{X})$ is equal to

$$\sup_{A \in \tilde{\mathcal{X}}} \phi(A) - \inf_{A \in \tilde{\mathcal{X}}} \phi(A). \tag{7.4}$$

The following facts are also easily verified. Sums and scalar multiples of measures on $\tilde{\mathcal{X}}$ correspond to sums and scalar multiples of the corresponding elements of $\mathfrak{C}^*(\tilde{X})$. We therefore, by an abuse of language, refer to measures on $\tilde{\mathcal{X}}$ as actually being linear functionals on $\mathfrak{C}(\tilde{X})$.

The set $\tilde{\mathbf{S}}$ of symmetric probabilities on $\tilde{\mathcal{X}}$ is convex, weakly closed, and bounded, considered as a subset of $\mathfrak{C}^*(\tilde{X})$. The extreme points of $\tilde{\mathbf{S}}$ are, as Theorem 5.3 shows, the elements of $\tilde{\boldsymbol{P}}$. Furthermore, the set $\tilde{\boldsymbol{P}}$ is weakly closed in $\mathfrak{C}^*(\tilde{X})$. We leave the verification of these assertions to the reader.

We now apply Theorem 6.3 with $L = \mathfrak{C}(\tilde{X})$, $L^* = \mathfrak{C}^*(\tilde{X})$, $N = \tilde{\mathbf{S}}$, and $P = \tilde{\boldsymbol{P}}$. This theorem implies that there exists a countably additive probability μ on the σ-algebra of weak Baire sets of $\tilde{\boldsymbol{P}}$ with the property that

$$(f, \sigma) = \int_{\tilde{\boldsymbol{P}}} (f, \tilde{\pi}) d\mu(\tilde{\pi}) \tag{7.5}$$

for all $f \in \mathfrak{C}(\tilde{X})$. Combining (7.3) and (7.5), we can write

$$\int_{\tilde{X}} f(a) d\sigma(a) = \int_{\tilde{\boldsymbol{P}}} \left[\int_{\tilde{X}} f(a) d\tilde{\pi}(a) \right] d\mu(\tilde{\pi}). \tag{7.6}$$

Relation (7.6) implies that

$$\sigma A = \int_{\tilde{\mathcal{X}}} (\tilde{\pi} A) d\mu(\tilde{\pi}) \tag{7.7}$$

for all $A \in \tilde{\mathcal{X}}$. To see this, we argue as follows. For every $B \subset \tilde{X}$ that is a closed G_δ, it is well known and obvious that there exists a descending sequence $\{f_n\}_{n=1}^{\infty}$ of functions in $\mathfrak{C}(\tilde{X})$ such that $\lim_{n \to \infty} f_n(a) = \chi_B(a)$ for all $a \in \tilde{X}$, where χ_B denotes as usual the characteristic function of the set B. Applying Lebesgue's convergence theorem, we then have

$$\sigma B = \int_{\tilde{X}} \chi_B(a) d\sigma(a) = \int_{\tilde{X}} \lim_{n \to \infty} f_n(a) d\sigma(a) = \lim_{n \to \infty} \int_{\tilde{X}} f_n(a) d\sigma(a).$$

By (7.6), the last expression is equal to

$$\lim_{n\to\infty} \int_{\tilde{P}} \left[\int_{\tilde{X}} f_n(a) d\tilde{\pi}(a) \right] d\mu(\tilde{\pi}). \tag{7.8}$$

Applying Lebesgue's convergence theorem twice to (7.8), we find that (7.8) is equal to

$$\begin{aligned} \int_{\tilde{P}} \left[\lim_{n\to\infty} \int_{\tilde{X}} f_n(a) d\tilde{\pi}(a) \right] d\mu(\tilde{\pi}) &= \int_{\tilde{P}} \left[\int_{\tilde{X}} \chi_B(a) d\tilde{\pi}(a) \right] d\mu(\tilde{\pi}) \\ &= \int_{\tilde{P}} (\tilde{\pi}B) d\mu(\tilde{\pi}). \end{aligned} \tag{7.9}$$

Note that if σ or μ failed to be countably additive, the application of Lebesgue's convergence theorem would be unjustified. Now, the measure σ is completely determined by its behavior on closed G_δ's [14, p. 229, Theorem H]. It follows that (7.7) is valid for all $A \in \tilde{\mathfrak{X}}$.

The foregoing argument establishes incidentally that $\tilde{\mathfrak{P}}^*$ is a sub-σ-algebra of the weak Baire sets in $\tilde{P}$. (It is also true that every weak Baire set in $\tilde{P}$ is in $\tilde{\mathfrak{P}}^*$; we omit the proof of this fact, which is not used in this paper.) Hence (7.7) holds if the probability μ is restricted to the σ-algebra $\tilde{\mathfrak{P}}^*$, as only sets in $\tilde{\mathfrak{P}}^*$ are needed to integrate the functions $\tilde{\pi}A$. We have therefore proved the following fundamental theorem.

THEOREM 7.2. *Let X be a compact Hausdorff space, and let $\mathfrak{X}$ be the σ-algebra of Baire sets in X. Then $\mathfrak{X}$ is presentable.*

We leave to the reader the proof of the following simple theorem.

THEOREM 7.3. *Let $\mathfrak{Y}$ be a presentable σ-algebra of subsets of a set Y. Let X be any nonvoid set in $\mathfrak{Y}$, and let $\mathfrak{X}$ be defined as the family of all subsets of X that are in $\mathfrak{Y}$. Then $\mathfrak{X}$ is a presentable σ-algebra of subsets of X.*

The Baire sets of a topological space are the elements of the σ-ring generated by compact G_δ's. The elements of the σ-algebra generated by compact G_δ's—that is, Baire sets and their complements—will be called here *wide* Baire sets. If X is a σ-compact and locally compact Hausdorff space, it is easily seen that every wide Baire set is a Baire set.

THEOREM 7.4. *Let Y be a locally compact Hausdorff space, and let X be a wide Baire subset of Y. Let $\mathfrak{X}$ consist of all wide Baire sets of Y that are contained in X. Then $\mathfrak{X}$ is presentable.*

Proof. In view of Theorem 7.3, it will be enough to prove the present theorem in the case $X = Y$.

Let αX denote the 1-point compactification of X (see for example [1, p. 93, Satz XIV]), and let q be the point adjoined to X in order to obtain αX. We distinguish 2 cases.

CASE I. X is σ-compact. In this case, it follows immediately from the definition of αX that q is a G_δ in αX and is accordingly a Baire set. Therefore $\{q\}' = X$ is also a Baire set. The present theorem follows in this case from Theorems 7.2 and 7.3.

CASE II. X is not σ-compact. In this case, q is plainly not a G_δ, and hence, as is known [14, p. 221, Theorem D], $\{q\}$ is not a Baire set. Therefore X is not a Baire set in αX. However, the Baire sets in X are intimately connected with the Baire sets in αX, as we now show. Let F be a closed G_δ in αX and let $f \in \mathfrak{C}(\alpha X)$ have the property that $F = \{y \mid y \in \alpha X,\ f(y) = 0\}$. The set $A = \{x \mid x \in X,\ f(x) \neq f(q)\}$ is obviously the union of a sequence of compact G_δ's, and hence A is a Baire set of X. If $f(q) \neq 0$, then F is contained in the Baire set A and hence F is a Baire set of X. If $q \in F$, then F' is contained in A, and F' is a Baire set of X. It follows that if B is a Baire set in αX, then $B \cap X$ is a wide Baire set of X. Conversely, if D is a Baire set in X, then D is a Baire set in αX, and if D' is a Baire set in X, then $D \cup \{q\}$ is a Baire set in αX. Thus the mapping $B \to B \cap X$ is a 1-to-1 mapping of the Baire sets in αX onto the wide Baire sets in X, that obviously preserves all countable Boolean operations. The presentability of $\mathfrak{X}$ now follows from the known presentability of the σ-algebra of Baire sets in αX (Theorem 7.2). This completes the present proof.

As an illustration of the preceding theorem, we observe that the family of Borel subsets of an arbitrary Borel set on the real line is a presentable σ-algebra. Since de Finetti [11, pp. 37–44] has proved the presentability of the σ-algebra of Borel sets on the real line, and since Theorem 7.3 is easy, this illustration has but little novelty. However, it is an important and far-reaching example since, so far as we know, every σ-algebra of sets known to have importance in applied science is σ-isomorphic to the σ-algebra of Borel subsets of some Borel set on the real line.

Of course, there exist compact Hausdorff spaces that do not have this property. A noncountable Cartesian product of 2-point spaces is one typical, and well known, example. Indeed, the Baire σ-algebra of this space is easily seen to be nonseparable with respect to the usual product measure, whereas the Borel σ-algebra of a Borel set is separable with respect to any measure [14, Theorem B, p. 168].

An easier but less interesting example is obviously provided by any compact Hausdorff space with more than exp ($\aleph_0$) Baire sets, say the Cartesian product of more than exp ($\aleph_0$) 2-point discrete spaces or the 1-point compactification of a discrete set of more than exp ($\aleph_0$) elements. Indeed, the σ-algebra of Baire sets of such a space is not even isomorphic to *any* σ-algebra contained in the class of Borel sets of the real line.

For still another example, consider the space $T_{\Omega+1}$ consisting of all ordinal numbers $\alpha \leqq \Omega$, Ω denoting as usual the smallest uncountable ordinal, with the ordinary order topology. It is easy to show that $A \subset T_{\Omega+1}$ is a Baire set if

and only if A is countable and excludes Ω or A' has these properties (see [16, p. 171, Remark 1]). For this σ-algebra, we have the following result.

THEOREM 7.5. *The σ-algebra of Baire sets in $T_{\Omega+1}$ is not isomorphic to the σ-algebra of Baire (Borel) subsets of a Baire (Borel) subset of any locally compact and σ-compact Hausdorff space Y satisfying Hausdorff's first axiom of countability.*

Proof. Assume that there exists a σ-isomorphism τ carrying the σ-algebra $\{A\}$ of Baire sets in $T_{\Omega+1}$ onto a σ-algebra $\{\tau A\}$ that consists of all Baire subsets of the Baire set $\tau T_{\Omega+1}=B\subset Y$. For each $\alpha<\Omega$ in $T_{\Omega+1}$, $\tau\{\alpha\}$ must be a single point in B. It is easy to see that B consists of at least $\aleph_1$ points and that every set τA is countable or has countable complement in B. By imbedding Y in its 1-point compactification (which also satisfies the first countability axiom, since Y is σ-compact), we may suppose that Y is actually compact. This being so, the set B admits at least 1 complete limit point p [2, p. 8, Théorème I]. Thus $U(p)\cap B$ is uncountable for every open neighborhood $U(p)$ of p. Assume that $U'(p)\cap B$ is countable for every such neighborhood. Then if $\{U_n(p)\}_{n=1}^{\infty}$ is a countable family of neighborhoods of p for which $\bigcap_{n=1}^{\infty} U_n(p)=\{p\}$, it follows that $(\bigcup_{n=1}^{\infty}U_n'(p))\cap B=\{p\}'\cap B$ is countable, and this is an obvious contradiction. It follows that B has at least 2 complete limit points, p_0 and p_1. Since there is a continuous real-valued function f on Y such that $f(p_0)=0$ and $f(p_1)=1$, the Baire subsets of B must include uncountable sets with uncountable complements. Thus the σ-algebra $\{A\}$ is not σ-isomorphic to the σ-algebra of Baire subsets of B. The same proof shows that $\{A\}$ is not σ-isomorphic to a σ-algebra of Borel sets in Y.

The foregoing theorem shows, in particular, that $\{A\}$ cannot be realized as the σ-algebra of Borel subsets of a Borel set on the real line.

Theorem 7.4 is, so far as we can tell, very general. It applies to all those σ-algebras of sets that Bourbaki in his recent monograph on integration [5] has singled out for detailed study. We cannot even show that there exists an $\mathfrak{X}$ to which the hypotheses of Theorem 7.4 do not apply. The apparent domain of applicability of Theorem 7.4 is somewhat enlarged by the following theorem.

THEOREM 7.6. *Let X be a completely regular space that is the union of a countable family of compact G_δ's, and let $\mathfrak{X}$ be all of the Baire sets of X. Then $\mathfrak{X}$ is σ-isomorphic to the σ-algebra of all wide Baire sets of a certain locally compact Hausdorff space.*

Proof. We sketch the proof. Let $X_1\subset X_2\subset\cdots\subset X_n\subset\cdots$ be compact G_δ's with union X. It is easy to see that each $X_n=\{x\,|\,x\in X, f_n(x)=0\}$ for some $f_n\in\mathfrak{C}(X)$. Let $\mathfrak{T}$ be the smallest family of subsets of X containing all sets open in the original topology of X and all sets $X_{n+1}\cap X_n'$ (set $X_0=0$), and

closed under the formation of arbitrary unions and finite intersections. It is easy to see that $\mathfrak{T}$ is the family of open subsets of X under a locally compact Hausdorff topology in which the wide Baire sets are just the wide Baire sets of the original topology on X.

As an illustration of the foregoing theorem, we note that the adjoint L^* of a Banach space L is evidently weakly completely regular, and it has the property that the sets $\{g \mid g \in L^*, \|g\| \leqq \lambda\}$ $(0 < \lambda < \infty)$ are weakly compact G_δ's if and only if L is separable.

8. **Presentability in the finite situation.** This section refers exclusively to the finite situation, except where the contrary is explicitly stated.

THEOREM 8.1. *Every Boolean algebra $\mathfrak{X}$ of subsets of a set X is presentable.*

Proof. This theorem can be proved by paraphrasing the proof of Theorem 7.2, and making use of the fact that the class of finitely additive real-valued measures on $\mathfrak{X}$ of finite total variation is a representation of the conjugate space of the normed linear space of all bounded $\mathfrak{X}$-measurable functions on X. It is, however, also interesting to prove the present theorem by a direct appeal to Theorem 7.2, as follows.

According to a representation theorem of M. H. Stone, $\mathfrak{X}$ is Boolean isomorphic to the algebra of compact open subsets in a certain totally disconnected compact Hausdorff space [25, p. 378, Theorem 1]. We may therefore suppose without loss of generality that X is a compact totally disconnected Hausdorff space and that $\mathfrak{X}$ is the algebra of all compact open subsets of X. $\tilde{X}$ is then a totally disconnected compact Hausdorff space, as can be easily seen, and the algebra $\tilde{\mathfrak{X}}$, which consists of all finite unions of cylinders $C(E_1, E_2, \cdots, E_n)$ with $E_i \in \mathfrak{X}$ $(i = 1, 2, \cdots, n)$, is exactly the algebra of all compact open subsets of $\tilde{X}$.

We now write $(\tilde{\mathfrak{X}})^+$ for the σ-algebra of subsets of $\tilde{X}$ generated by $\tilde{\mathfrak{X}}$. Since every continuous real-valued function on $\tilde{X}$ is arbitrarily uniformly approximable by linear combinations of characteristic functions of sets in $\tilde{\mathfrak{X}}$, it follows that $(\tilde{\mathfrak{X}})^+$ is the σ-algebra of Baire sets of $\tilde{X}$.

Now let $\mathfrak{Y}$ denote the σ-algebra of subsets of X generated by the algebra $\mathfrak{X}$, and let $\tilde{\mathfrak{Y}}$ denote the σ-algebra of subsets of $\tilde{X}$ generated by cylinders $C(E_1, E_2, \cdots, E_n)$ with $E_i \in \mathfrak{Y}$ $(i = 1, 2, \cdots, n)$. It is easy to see that $\tilde{\mathfrak{Y}} = (\tilde{\mathfrak{X}})^+$.

Now let σ be any finitely additive symmetric probability on $\tilde{\mathfrak{X}}$. It is obvious that σ is countably additive, since if $A_1 \supset A_2 \supset A_3 \supset \cdots \supset A_n \supset \cdots$ is any descending sequence of sets in $\tilde{\mathfrak{X}}$ with void intersection, some A_n is already void, since the A_n's are all compact subsets of X. (See also [28, §4]). A known extension theorem (stated, for example, in [14, p. 54, Theorem A]) then implies that σ admits a unique extension σ^+ over the σ-algebra $(\tilde{\mathfrak{X}})^+$. By Theorem 7.2, we know that there exists a probability μ^+ on the σ-algebra

$\tilde{\mathcal{P}}^{+*}(\tilde{P}^{+}=$all product probabilities on $\tilde{\Upsilon}=(\tilde{\mathcal{X}})^{+})$ such that

$$\sigma^{+}A = \int_{P^{+}} (\pi^{+}A)d\mu^{+}(\pi^{+}), \tag{8.1}$$

for all $A\in(\mathcal{X})^{+}$. In particular, upon contracting σ^{+} and π^{+} to the algebra $\tilde{\mathcal{X}}$, we can re-write (8.1) as

$$\sigma B = \int_{P^{+}} (\pi B)d\mu^{+}(\pi), \tag{8.2}$$

for all $B\in\tilde{\mathcal{X}}$.

Every finitely additive product probability π on $\tilde{\mathcal{X}}$ is countably additive and thus, like the more general σ's considered above, admits a unique extension π^{+} over $(\tilde{\mathcal{X}})^{+}$. The sets P and P^{+} can thus be identified, and the probability μ^{+} referred to in (8.1) and (8.2) can be identified with a probability μ on the σ-algebra of subsets of P generated by all sets of the form

$$\{\pi \mid \pi \in P,\ \pi A \leqq \lambda\} \qquad (0 \leqq \lambda \leqq 1) \tag{8.3}$$

for $A\in(\tilde{\mathcal{X}})^{+}$. This probability μ can obviously be contracted to the σ-algebra generated by all sets of the form

$$\{\pi \mid \pi \in P,\ \pi B \leqq \lambda\} \qquad (0 \leqq \lambda \leqq 1) \tag{8.4}$$

for $B\in\tilde{\mathcal{X}}$. Formula (8.2) can thus be re-written as

$$\sigma B = \int_{P} (\pi B)d\mu(\pi) \tag{8.5}$$

for all $B\in\tilde{\mathcal{X}}$. Therefore $\mathcal{X}$ is presentable.

9. **Uniqueness of μ for presentable σ.** For technical simplicity, this section and §10 are confined to the infinite situation. A minor complication in the finite situation is that we here compute a symmetric probability σ for certain sets that need not be included in $\tilde{\mathcal{X}}$ in the finite situation. However, it is easy to show that σ has a unique natural extension to these sets and that the theorems of the present section and §10 can be paralleled in the finite situation. This can be done either by elementary techniques used by de Finetti [11, pp. 31–32] or by using Stone's representation theorem as in the proof of Theorem 8.1. It is interesting to compare this section with §3 of [7].

Let σ be a presentable symmetric probability on $\tilde{\mathcal{X}}$. Then, by definition, σ admits at least one representation in the form

$$\sigma A = \int_{P} (\tilde{\pi}A)d\mu(\pi), \tag{9.1}$$

for all $A \in \tilde{\mathcal{X}}$, where μ is a countably additive probability on the σ-algebra $\mathcal{P}^*$. We shall give an algorithm for computing μ in terms of (9.1).

With $N(E, \lambda)$ defined as in (2.2), we introduce the further abbreviation

$$(9.2)\quad N(E_1, \cdots, E_n; \alpha_1, \cdots, \alpha_n; \beta_1, \cdots, \beta_n) = \bigcap_{i=1}^{n} N(E_i, \alpha_i) \cap \bigcap_{i=1}^{n} N(E_i, \beta_i)'.$$

Next, for all $a = \{a_n\}_{n=1}^{\infty} \in \tilde{X}$, $E \in \mathcal{X}$, and $n = 1, 2, 3, \cdots$, we define

$$\phi(a; E, n) = \frac{1}{n} \sum_{j=1}^{n} \chi_E(a_j).$$

THEOREM 9.1. *Let $E \in \mathcal{X}$ and let σ be a not necessarily presentable symmetric probability on $\tilde{\mathcal{X}}$. Then*

$$(9.3)\quad \lim_{n \to \infty} \phi(a; E, n)$$

exists for all $a \in \tilde{X}$, except perhaps for a set of σ-measure 0.

Proof. Let $\mathcal{X}_0$ be the algebra $\{0, E, E', X\}$. Theorem 7.2 implies that $\mathcal{X}_0$ is presentable. Let A be the subset of $\tilde{X}$ for which (9.3) exists. It is clear that $A \in \tilde{\mathcal{X}}_0$. We next observe that for all $\tilde{\pi}_0 \in \tilde{P}_0$, we have $\tilde{\pi}_0 A = 1$. This follows from the most elementary strong law of large numbers. (See for example [14, p. 205, ex. 7].) On the σ-algebra $\tilde{\mathcal{X}}_0$, the symmetric probability σ is presentable, and thus by the analogue of (9.1) that applies to σ on this σ-algebra, we infer that $\sigma A = 1$, as was to be proved.

It has been pointed out to us that the preceding theorem is an immediate consequence of a more general but deeper law of large numbers, which holds for all strictly stationary stochastic processes with discrete parameter [8, p. 465, Theorem 2.1].

We next define a certain $\tilde{\mathcal{X}}$-measurable function $\phi(a; E)$, by the relations

$$(9.4)\quad \phi(a; E) = \begin{cases} \lim_{n \to \infty} \phi(a; E, n) \text{ where this limit exists,} \\ 0 \text{ elsewhere.} \end{cases}$$

Furthermore, let

$$(9.5)\quad M(E; a) = \{a \mid a \in \tilde{X}, \phi(a; E) \leqq \alpha\} \qquad (0 \leqq \alpha \leqq 1).$$

Also let

$$(9.6)\quad M(E_1, \cdots, E_n; \alpha_1, \cdots, \alpha_n; \beta_1, \cdots, \beta_n) = \bigcap_{i=1}^{n} M(E_i; \alpha_i) \cap \bigcap_{i=1}^{n} M(E_i; \beta_i)'$$

for all $E_i \in \mathcal{X}$ and real numbers α_i and β_i $(i = 1, \cdots, n)$.

THEOREM 9.2. *Suppose that σ and μ are connected by the relation* (9.1). *Then*

$$\begin{aligned}(9.7)\qquad &\mu N(E_1, \cdots, E_n; \alpha_1, \cdots, \alpha_n; \beta_1, \cdots, \beta_n)\\ &\qquad = \sigma M(E_1, \cdots, E_n; \alpha_1, \cdots, \alpha_n; \beta_1, \cdots, \beta_n)\end{aligned}$$

for all $E_i \in \mathcal{X}$ *and real numbers* α_i *and* β_i $(i=1, \cdots, n)$.

Proof. The strong law of large numbers, in the form used in proving Theorem 9.1, asserts that

$$\phi(a; E) = \pi E \tag{9.8}$$

almost everywhere on X with respect to π. It follows that

$$\tilde{\pi} M = \chi_N(\pi), \tag{9.9}$$

where M and N are obvious abbreviations for the sets written in (9.7). Therefore we have

$$\begin{aligned}(9.10)\qquad \sigma M &= \int_{\tilde{P}} (\tilde{\pi} M) d\mu(\tilde{\pi})\\ &= \int_{\tilde{P}} \chi_N(\tilde{\pi}) d\mu(\tilde{\pi})\\ &= \mu N,\end{aligned}$$

as was to be proved.

Now, let $\mathcal{N}^*$ denote the family of subsets of $\mathcal{P}^*$ of the form (9.2).

THEOREM 9.3. $\mathcal{N}^*$ *is a semialgebra that generates the* σ*-algebra* $\mathcal{P}^*$.

Proof. It is obvious from (9.2) that $\mathcal{N}^*$ is closed under the formation of finite intersections. The set P itself can be written as $N(E; 1; -1)$ for arbitary $E \in \mathcal{X}$. Finally,

$$\begin{aligned}(9.11)\qquad &N(E_1, \cdots, E_n; \alpha_1, \cdots, \alpha_n; \beta_1, \cdots, \beta_n)'\\ &\qquad = N(E_1, \cdots, E_{n-1}; \alpha_1, \cdots, \alpha_{n-1}; \beta_1, \cdots, \beta_{n-1})'\\ &\qquad \cup N(E_n; \beta_n; -1) \cup N(E_n; 1; \alpha_n).\end{aligned}$$

An obvious induction and the fact that no generality is lost in assuming $\beta_n \leqq \alpha_n$ now suffice to show that $\mathcal{N}^*$ is a semialgebra.

It is obvious that $\mathcal{N}^*$ generates the σ-algebra $\mathcal{P}^*$.

THEOREM 9.4. *Let* σ *be a symmetric probability on* $\tilde{\mathcal{X}}$ *and let* μ *and* μ' *both stand in the relation* (9.1) *to* σ. *Then* $\mu = \mu'$ *for all sets in* $\mathcal{P}^*$.

Proof. Theorem 9.2 shows that $\mu N = \mu' N$ for all $N \in \mathcal{N}^*$. Theorems 3.1 and 9.3 imply that $\mu A = \mu' A$ for all $A \in \mathcal{P}^*$.

10. **An algorithmic study of presentability.** When σ is presentable, the corresponding probability μ is determined by the algorithm summarized in (9.7). This suggests that we study the presentability of a given $\sigma \in \tilde{\mathbf{S}}$ by trying

to see whether or not defining μ as in (9.7) leads to a probability on $\mathcal{P}^*$ for which (9.1) is satisfied. This program, which is much more constructive than that utilized in the proof of Theorem 7.2, might be expected to be very powerful. However, we have had only limited success with it, which culminates in new formulations of the condition of presentability (Theorem 10.5). We note also that the line of reasoning of this section is essentially that adopted by Dynkin [9], who with it is able to prove the presentability of the σ-algebra of Borel sets on the real line and of other σ-algebras that are in a certain sense separable.

This section is, for simplicity, restricted to the infinite situation, as was explained in the first paragraph of §9.

We shall call a triple of arguments $\{E_1, \cdots, E_n; \alpha_1, \cdots, \alpha_n; \beta_1, \cdots, \beta_n\}$ ($E_i \in \mathcal{X}$, α_i, β_i real numbers, $1 \leqq i \leqq n$) a *tile*, and we shall call the sets E_i its *constituents*.

THEOREM 10.1. *Let* $\{T_u\}$, $\{T_v^*\}$ *be finite sequences of tiles; let* $\mathcal{X}_0$ *be a finite subalgebra of* $\mathcal{X}$ *containing all the constituents of these tiles; let*

$$B = \{a \mid a \in \tilde{X}, \lim_{n\to\infty} \phi(a, E; n) \text{ exists for all } E \in \mathcal{X}_0\}; \tag{10.1}$$

and let

$$\boldsymbol{N} = \bigcup_u \boldsymbol{N}(T_u) \Delta \bigcup_v \boldsymbol{N}(T_v^*), \qquad M = \bigcup_u M(T_u) \Delta \bigcup_v M(T_v^*), \tag{10.2}$$

where "Δ" denotes symmetric difference. The following conditions are logically equivalent:

(a) $M \cap B = 0$;
(b) $\sigma M = 0$ *for all* $\sigma \in \tilde{\mathbf{S}}$;
(c) $\tilde{\pi} M = 0$ *for all* $\tilde{\pi} \in \tilde{\boldsymbol{P}}$;
(d) $\boldsymbol{N} = 0$.

Proof. Theorem 9.1 implies that $\sigma B = 1$, since $\mathcal{X}_0$ is finite. Hence $\sigma M = \sigma(M \cap B)$, so that (a) implies (b).

It is trivial that (b) implies (c).

The fact that (c) implies (d) follows directly from (9.8); if $\pi \in \boldsymbol{N}$, then $\tilde{\pi} M = 1$.

To show that (d) implies (a), let a be a point of $\tilde{X}$ in $M \cap B$. The algebra $\mathcal{X}_0$ is generated by the elements $F_1, \cdots, F_k$ of a finite partition of X. Let $e_i \in F_i$ ($i = 1, \cdots, k$) and let π be the set-function defined for all $E \in \mathcal{X}$ by the relation

$$\pi E = \sum_{i=1}^{k} \phi(a; F_i) \chi_E(e_i). \tag{10.3}$$

It is easy to show that π is an element of $\boldsymbol{N}$. This completes the proof.

Now let σ be an arbitrary element of $\tilde{\mathbf{S}}$, and let $\tilde{\mathcal{X}}_\sigma$ be the Boolean σ-algebra

obtained by reducing $\tilde{\mathfrak{X}}$ modulo sets of σ-measure 0. For $A\in\tilde{\mathfrak{X}}$, let $[A]_\sigma$ be the element of $\tilde{\mathfrak{X}}_\sigma$ containing A. The following two theorems are simple consequences of Theorems 9.3, 10.1, and 3.1. We leave the proofs to the reader.

THEOREM 10.2. *There is a unique Boolean homomorphism of* $\mathcal{B}^*(\mathcal{N}^*)$, *the Boolean algebra generated by* $\mathcal{N}^*$, *into* $\tilde{\mathfrak{X}}_\sigma$ *that carries* $\boldsymbol{N}(T)$ *onto* $[M(T)]_\sigma$ *for every tile* T.

THEOREM 10.3. *For every* $\sigma\in\tilde{\mathbf{S}}$, *there is a unique finitely additive probability* μ *on* $\mathcal{B}^*(\mathcal{N}^*)$ *such that* $\mu\boldsymbol{N}(T)=\sigma M(T)$ *for every tile* T.

THEOREM 10.4. *If* σ, $\sigma'\in\tilde{\mathbf{S}}$ *and* $\sigma M(T)=\sigma' M(T)$ *for every tile* T, *then* $\sigma=\sigma'$.

Proof. First, if both σ and σ' are presentable, the present theorem follows at once from (9.1) and Theorem 9.4. For general σ and σ', consider an arbitrary cylinder $C(E_1, \cdots, E_n)\in\tilde{\mathfrak{X}}$ (i.e., $E_1, \cdots, E_n\in\mathfrak{X}$). Let $\mathfrak{X}_0$ be a finite (and hence presentable) subalgebra of $\mathfrak{X}$ that contains $E_1, \cdots, E_n$, and let $\tilde{\mathfrak{X}}_0$ be the corresponding subalgebra of $\tilde{\mathfrak{X}}$. Considering the contractions of σ and σ' to the algebra $\tilde{\mathfrak{X}}_0$, we see that $\sigma A=\sigma' A$ for all $A\in\tilde{\mathfrak{X}}_0$. Therefore, $\sigma C(E_1, \cdots, E_n)=\sigma' C(E_1, \cdots, E_n)$. It follows from Theorem 3.1 that $\sigma A=\sigma' A$ for all $A\in\tilde{\mathfrak{X}}$.

THEOREM 10.5. *For every* $\sigma\in\tilde{\mathbf{S}}$, *the following* 4 *conditions are logically equivalent*:

(a) σ *is presentable*;

(b) *if* $A\in\tilde{\mathfrak{X}}$ *and* $\sigma A\neq 0$, *there exists* $\tilde{\pi}\in\tilde{\boldsymbol{P}}$ *such that* $\tilde{\pi}A\neq 0$;

(c) *the homomorphism described in Theorem* 10.2 *carries sequences descending to* 0 *into sequences with the same property*;

(d) *the function* μ *described in Theorem* 10.3 *is countably additive.*

Proof. It is clear from (9.1) that (a) implies (b).

To show that (b) implies (c), let $\{\boldsymbol{N}_i\}_{i=1}^{\infty}$ be a sequence in $\mathcal{B}^*(\mathcal{N}^*)$ descending to 0. The image of $\{\boldsymbol{N}_i\}_{i=1}^{\infty}$ in $\tilde{\mathfrak{X}}_\sigma$ is clearly descending; suppose that it does not descend to 0. Represent $\boldsymbol{N}_i$ in the form of a disjoint union of elements of $\mathcal{N}^*$:

$$\boldsymbol{N}_i=\bigcup_{j=1}^{n}\boldsymbol{N}(T_{j,i}). \tag{10.4}$$

The image of $\boldsymbol{N}_i$ in $\tilde{\mathfrak{X}}_\sigma$ is then $[M_i]_\sigma$, where

$$M_i=\bigcup_{j=1}^{n}M(T_{j,i}). \tag{10.5}$$

By hypothesis, $[M_i]_\sigma$ does not descend to 0. Hence, for all i, $\sigma M_i\neq 0$, and, if (b) holds, there exists a probability $\pi_0\in\boldsymbol{P}$ such that $\tilde{\pi}_0 M_i\neq 0$. It follows im-

mediately from this and from (9.9) that $\tilde{\pi}_0 \in \boldsymbol{N}_i$ for all i. This contradiction proves that (b) implies (c).

Suppose next that (c) holds, and that $\{\boldsymbol{N}_i\}_{i=1}^{\infty}$ is a sequence in $\mathcal{B}^*(\mathcal{N}^*)$ that descends to 0. Now, in the notation introduced in (10.4) and (10.5), we have $\mu\boldsymbol{N}_i = \sigma M_i$, and it follows that $\lim_{i\to\infty} \mu\boldsymbol{N}_i = 0$ since σ is countably additive. Hence μ is countably additive, by a well-known criterion, and this is just condition (d).

Finally, suppose that (d) holds. Then the probability μ admits a (unique) countably additive extension over the σ-algebra $\mathcal{P}^*$. We shall show that (9.1) holds for this μ and our probability σ, and this will establish (a). Consider the set-function σ' defined for all $A \in \tilde{\mathcal{X}}$ by the relation

$$\sigma' A = \int_{\tilde{P}} (\tilde{\pi} A) d\mu(\tilde{\pi}). \tag{10.6}$$

It is obvious that σ' is a presentable element of $\tilde{\mathbf{S}}$. For every tile T, we find, with the aid of (9.9), that

$$\begin{aligned} \sigma' M(T) &= \int_{\tilde{P}} (\tilde{\pi} M(T)) d\mu(\tilde{\pi}) \\ &= \mu \boldsymbol{N}(T) \\ &= \sigma M(T). \end{aligned} \tag{10.7}$$

Theorem 10.4 implies that $\sigma = \sigma'$; hence σ is presentable. This is condition (a), and thus the present proof is complete.

11. **A generalization.** The study of presentability carried out above is the study of a set $\tilde{\mathbf{S}}$ of probabilities invariant under a certain group G of transformations of the underlying space, with special reference to the extreme points of $\tilde{\mathbf{S}}$. In the present section, we consider a somewhat different situation from that treated above, in which $\tilde{X}$, $\tilde{\mathcal{X}}$, and the relevant group of transformations are all changed. Though we have not carried this different investigation to the point of generalizing the theorems of §7, it provides a convenient avenue for studying finite Cartesian products (§12). In the present section, we assign new meanings to familiar notation and terminology, in order to emphasize analogies. For simplicity's sake, we confine our attention to the infinite situation.

Throughout the present section, let $\tilde{X}$ denote a set and $\tilde{\mathcal{X}}$ a σ-algebra of subsets of $\tilde{X}$. Let G be a class of σ-isomorphisms of $\tilde{\mathcal{X}}$ into itself, $\tilde{\mathbf{S}}$ the set of all probabilities on $\tilde{\mathcal{X}}$ that are invariant under all transformations of G, and $\tilde{\boldsymbol{P}}$ the set of extreme points of $\tilde{\mathbf{S}}$. (We regard $\tilde{\mathbf{S}}$ as a convex set in the set of all measures on $\tilde{\mathcal{X}}$, in the usual way.)

A probability $\sigma \in \tilde{\mathbf{S}}$ is said to be presentable if there exists a probability μ on the σ-algebra of subsets $\mathcal{P}^*$ of $\tilde{\boldsymbol{P}}$ generated by sets of the form $\{\tilde{\pi} \mid \tilde{\pi} \in \tilde{\boldsymbol{P}}, \tilde{\pi} A \leqq \lambda\}$ for λ real and $A \in \tilde{\mathcal{X}}$ such that

$$\sigma A = \int_{\mathbf{P}} (\tilde{\pi} A) d\mu(\tilde{\pi}) \tag{11.1}$$

for all $A \in \tilde{\mathfrak{X}}$. If all $\sigma \in \tilde{\mathbf{S}}$ are presentable, then $(\tilde{\mathfrak{X}}, G)$ is said to be presentable. We do not know whether or not every $(\tilde{\mathfrak{X}}, G)$ is presentable. The following paragraphs culminate in the characterization of $\tilde{\boldsymbol{P}}$ in case G is a finite group. The conclusions reached are mostly very simple and can hardly be new.

THEOREM 11.1. *Let X be a set and $\mathfrak{X}$ an algebra of subsets of X. If $\boldsymbol{M}$ is a convex set of probabilities on X, if $\mathfrak{Y} \subset \mathfrak{X}$ is closed under the formation of complements, and if, for every $\pi \in \boldsymbol{M}$ and $E \in \mathfrak{Y}$ for which $\pi E > 0$, the probability $\pi \mid E$ is in $\boldsymbol{M}$, then $\pi E = 0$ or 1 for every extreme point π of $\boldsymbol{M}$ and every $E \in \mathfrak{Y}$.*

THEOREM 11.2. *Let $\tilde{X}$, $\tilde{\mathfrak{X}}$, G, $\tilde{\mathbf{S}}$ and $\tilde{\boldsymbol{P}}$ be as defined above. Then, if $\tilde{\pi} \in \tilde{\boldsymbol{P}}$, $A \in \tilde{\mathfrak{X}}$, and A is invariant under all transformations in G, then $\tilde{\pi} A = 0$ or 1.*

The proofs of these two theorems are very simple, and are omitted.

The following theorem, which is in the nature of a digression, is of some interest in the theory of probability. It says somewhat more than the 0-or-1 law as applied to products of identical distributions, even in the simplest cases [14, p. 201, exercise 3].

THEOREM 11.3. *A product measure on an infinite product of measure spaces can assume only the values 0 and 1 for sets that are invariant under all finite permutations of the co-ordinates.*

Proof. This result follows immediately from Theorems 5.3 and 11.2.

The preceding theorem was commented on by several who saw a prepublication copy of this paper. Blackwell, and Chung and Derman wrote us independently that they had become interested in the following question in connection with forthcoming publications. Is it true that the partial sums of a sequence of identically distributed independent random variables visit an arbitrary Borel set infinitely often with probability either 0 or 1? As they point out, the affirmative answer, which they had already demonstrated in certain cases, is an immediate consequence of Theorem 11.3. Halmos and Doob have shown us direct proofs of Theorem 11.3, both of which make it plain that that theorem is close to and scarcely deeper than the ordinary 0-or-1 law. These proofs are, with their authors' permission, presented below.

Following Halmos, suppose the sequence of random variables to be indexed by $\cdots, -1, 0, 1, \cdots$. An invariant set is easily shown to be almost invariant under the shift transformation; it must therefore have probability 0 or 1, because this transformation is ergodic. Alternatively, as Doob has pointed out, an invariant set is easily shown to be independent of itself, and hence to have probability 0 or 1.

It is interesting to quote *verbatim* a proof by Doob in a rather different spirit.

"You might be interested in my approach to Theorem 11.3. I'll state it in the language of random variables: Let $x_1, x_2, \cdots$ be a sequence of mutually ind. r.v. with a common distribution, and let x be any r.v. dependent on them and invariant in the sense described. Then x is a constant (prob. 1). To prove this it is no restriction to assume that x is bounded and has expectation 0. Now consider the conditional expectation $E\{x|x_1\}$, that is, the projection of x_1 on the linear manifold of Baire functions of x_1 which are in L_2. By hypothesis this can be written in the form $\phi(x_1)$, where $E\{x|x_n\}=\phi(x_n)$ also. Now consider the sequence $\phi(x_1), \phi(x_2), \cdots$. These r.v. are independent, with 0 expectations, and are therefore orthogonal. By Bessel's inequality the sum of the squares of their L_2 norms is at most the square of the norm of x. Since all the norms are equal, they vanish, so $E\{x|x_1\}=0$, and in exactly the same way $E\{x|x_1, \cdots, x_n\}=0$ for all n. When $n\to\infty$ this conditional expectation becomes x, so $x=0$, as was to be proved. This proof is of course just a slightly different twist to Levy's proof of the usual 0-1 law, which is simpler because, by hypothesis, the last cond. exp. above is a constant, so that x must also be a constant."

THEOREM 11.4. *Let G be a finite group. Then an element $\sigma\in\tilde{\mathbf{S}}$ is in $\tilde{\mathbf{P}}$ if and only if $\sigma A=0$ or 1 for every $A\in\tilde{\mathfrak{X}}$ that is invariant under all transformations of G.*

Proof. The necessity of the condition stated was pointed out in Theorem 11.2.

To prove the sufficiency, we proceed as follows. Let $\sigma\in\tilde{\mathbf{S}}$ be a nonextreme point; then there exists a real number $\alpha\in]0, 1[$, probabilities $\sigma', \sigma''\in\tilde{\mathbf{S}}$, and $B\in\tilde{\mathfrak{X}}$ such that $\sigma' B\neq\sigma'' B$ and $\sigma=\alpha\sigma'+(1-\alpha)\sigma''$. Let $\mathcal{B}$ denote the algebra of subsets of $\tilde{X}$ generated by the images of B under all of the transformations of G. Since G is finite, it follows that $\mathcal{B}$ is finite and is hence generated by the elements $\{B_i\}_{i=1}^m$ of a finite partition of X. Since B itself is partitioned by the sets B_i $(i=1, \cdots, m)$, there exists an index i_0 such that $\sigma' B_{i_0}\neq\sigma'' B_{i_0}$. Every image of B_{i_0} under a transformation of G is an element of $\{B_i\}_{i=1}^m$. Let the union of these images be denoted by C. Then clearly $\sigma' C\neq\sigma'' C$, and C is invariant under all transformations of G. It follows that σC is neither 0 nor 1, and the theorem is proved.

The hypothesis in Theorem 11.4 that G be a finite group can be replaced by others. For example, if G is the integers or the semigroup of positive integers under addition, known ergodic theorems can be invoked [22]. However, Theorem 11.4 is not true for all groups G. For example, if $\tilde{\mathfrak{X}}$ consists of the Baire sets of $\tilde{X}$, where $\tilde{X}$ is the Cartesian product of an uncountable number of discrete spaces each containing exactly 2 points, and G is the group of *all* permutations of co-ordinates, Theorem 11.4 is false. Indeed it is easily verified that $\tilde{\mathbf{S}}$ has more than 1 element, but that they all take only the values 0 and 1 on all invariant Baire sets, namely the null set and $\tilde{X}$.

THEOREM 11.5. *Let G be a finite group, of order n, and let θ be a not necessarily symmetric probability on $\tilde{\mathfrak{X}}$. Then there exists a unique element σ_θ of $\tilde{\mathbf{S}}$ such that $\sigma_\theta A = \theta A$ for all $A \in \tilde{\mathfrak{X}}$ that are invariant under all transformations of G. This probability σ_θ is determined by the relation*

$$\sigma_\theta B = \frac{1}{n} \sum_{g \in G} \theta(gB) \tag{11.2}$$

for all $B \in \tilde{\mathfrak{X}}$.

Proof. It is clear that (11.2) determines an element of $\tilde{\mathbf{S}}$ with the required property. Its uniqueness is easily proved by the technique of Theorem 11.4.

If θ is unit mass concentrated at a point of X (a 1-point probability), it is intuitively plausible that the corresponding measure σ_θ lies in $\tilde{\mathbf{P}}$; the following theorem shows this and more. It is expressed in terms of 0-1 probabilities, that is, probabilities that assume only the values 0 and 1. These are, as is well known, more general than point probabilities. See for example [28].

THEOREM 11.6. *Let G be a finite group, of order n. Then the elements of $\tilde{\mathbf{P}}$ all have the form σ_θ for 0-1 probabilities θ. If θ is a 0-1 probability, then σ_θ lies in $\tilde{\mathbf{P}}$. For 0-1 probabilities θ and θ', we have $\sigma_\theta = \sigma_{\theta'}$ if and only if there exists $g_0 \in G$ such that $\theta' B = \theta(g_0 B)$ for all $B \in \tilde{\mathfrak{X}}$. (In this case, θ and θ' are said to be congruent under G.)*

Proof. Let $\tilde{\pi} \in \tilde{\mathbf{P}}$, and let p be the infimum of positive values assumed by π. It follows at once from Theorem 11.4 that $np \geqq 1$ and that there exists $A \in \tilde{\mathfrak{X}}$ such that $\tilde{\pi} A = p$. Then the conditional probability $\tilde{\pi} \mid A = \theta$ is a 0-1 measure, and $\sigma_\theta = \tilde{\pi}$. For, if we assume the contrary, it is easy to apply the technique used in proving Theorem 11.4 to show that there exists a $C \in \tilde{\mathfrak{X}}$ invariant under all transformations of G such that

$$\tilde{\pi} C \neq \sigma_\theta C = \frac{1}{pn} \sum_{g \in G} \tilde{\pi}(A \cap gC) = \frac{1}{p} \tilde{\pi}(A \cap C). \tag{11.3}$$

Since (11.3) is a contradiction for $\tilde{\pi} C = 0$ and also for $\tilde{\pi} C = 1$, and since C is invariant, Theorem 11.4 implies that σ_θ must be $\tilde{\pi}$.

The second statement of the present theorem is obvious in view of Theorem 11.4.

To prove the last statement, suppose that θ and θ' are 0-1 measures for which $\sigma_\theta = \sigma_{\theta'}$, and assume that there is no $g \in G$ for which $\theta' B = \theta(gB)$ for all $B \in \tilde{\mathfrak{X}}$. Then, for every $g \in G$, there exists a set $B_g \in \tilde{\mathfrak{X}}$ such that $\theta(B_g) = 0$ and $\theta'(B_g) = 1$. Let $C = \bigcap_{g \in G} B_g$. Then $\theta' C = 1$ and $\theta(gC) = 0$, for all $g \in G$. Now we have

$$n\sigma_\theta C = \sum_{g \in G} \theta(gC) = 0, \quad n\theta' C = \sum_{g \in G} \theta'(gC) > 0, \tag{11.4}$$

and this is a contradiction. The proof is complete.

THEOREM 11.7. *If G is a finite group, then $(\tilde{\mathfrak{X}}, G)$ is presentable, and for every $\sigma \in \tilde{\mathbf{S}}$, there is exactly* 1 *measure μ satisfying relation* (11.1). *This μ is determined by the condition that*

$$\mu\{\tilde{\omega} \mid \tilde{\pi}A = 1\} = \sigma A \tag{11.5}$$

for all $A \in \tilde{\mathfrak{X}}$ that are invariant under all transformations of G.

Proof. Theorem 11.2 shows that (11.1) implies (11.5) for all G, whether finite or infinite. The technique used in proving Theorem 11.4 can be applied to show at once that (11.6) can hold for at most 1 probability μ.

If G consists of the identity transformation alone, then, as Theorem 11.4 shows, $\tilde{\boldsymbol{P}}$ consists exactly of the 0-1 probabilities on $\tilde{\mathfrak{X}}$, and the integral representation (11.1) exists if and only if (11.5) obtains. Thus (11.1) holds if and only if the set-function μ defined in (11.5) has a countably additive extension to the Boolean algebra consisting of the sets $\{\tilde{\pi} \mid \tilde{\pi}A = 1\}$ for $A \in \tilde{\mathfrak{X}}$. Using point probabilities, it is easy to see that if $\{A_n\}_{n=1}^{\infty}$ is a sequence of sets in $\tilde{\mathfrak{X}}$ for which the sets $\{\tilde{\pi} \mid \tilde{\pi}A_n = 1\}$ are pairwise disjoint, then the sets A_n are pairwise disjoint. This proves that μ as defined in (11.5) is countably additive. Therefore $(\tilde{\mathfrak{X}}, G)$ is presentable if G consists of the identity alone.

To prove the general case, let ν be a probability on the set $\boldsymbol{R}$ of 0-1 measures θ such that

$$\sigma A = \int_{\boldsymbol{R}} \theta d\nu(\theta) \tag{11.6}$$

for all $A \in \tilde{\mathfrak{X}}$ and a given $\sigma \in \tilde{\mathbf{S}}$. The special case considered above shows that such a probability ν exists. Now let μ be the set-function on $\tilde{\mathcal{P}}^*$ such that, for all $\boldsymbol{\mathcal{Q}} \in \tilde{\mathcal{P}}^*$,

$$\mu\boldsymbol{\mathcal{Q}} = \nu\{\theta \mid \sigma_\theta \in \boldsymbol{\mathcal{Q}}\}. \tag{11.7}$$

It is easy to show that this μ is a countably additive probability on $\mathcal{P}^*$ for which (11.1) is satisfied, and thus the present proof is complete.

It is instructive to note that this section could easily have been formulated in terms of an abstract σ-algebra rather than in terms of the algebra of sets $\tilde{\mathfrak{X}}$. Virtually the whole section would have remained unchanged, with the crucial exception of Theorem 11.7. The proof of that theorem makes essential use of point probabilities, and it is in fact false for abstract σ-algebras. To see this it is enough to recall or demonstrate that there are no 0-1 measures on such a σ-algebra as the Borel sets modulo sets of Lebesgue measure 0 on the real line.

12. **The presentability of finite Cartesian products.** In the present section, we study the exact sense in which finite Cartesian products are pre-

sentable. Since the finite situation is relatively uninteresting here, we limit ourselves to the infinite situation.

Let X be a set, $\mathcal{X}$ an algebra of subsets of X, and n an integer >1. By $\tilde{X}_n$ we shall mean the Cartesian product of n replicas of X, and by $\tilde{\mathcal{X}}_n$ the smallest algebra containing all sets $C(E_1, \cdots, E_n) = E_1 \times \cdots \times E_n$ for $E_1, \cdots, E_n \in \mathcal{X}$. For $E \in \mathcal{X}$, we write $C(E)$ for $E \times X_2 \times \cdots \times X_n$. Let G denote the group of all permutations of co-ordinates on $\tilde{X}_n$. Let $\tilde{\mathbf{S}}_n$ be the convex set of all probabilities on $\tilde{\mathcal{X}}_n$ invariant under all the transformations of G, and let $\tilde{\boldsymbol{P}}_n$ be the set of extreme points of $\tilde{\mathbf{S}}_n$.

Theorem 11.7 implies that $(\tilde{\mathcal{X}}_n, G)$ is presentable in the sense that a representation of the form (11.1) exists for every $\sigma \in \tilde{\mathbf{S}}_n$. This representation is unique, and furthermore, the elements of $\tilde{\boldsymbol{P}}_n$ are simply the elements of the form σ_θ for 0-1 probabilities θ on $\tilde{\mathcal{X}}_n$.

Our first task is to characterize the elements of $\tilde{\boldsymbol{P}}_n$ in terms of $\mathcal{X}$ rather than of $\tilde{\mathcal{X}}_n$. To do this, we define $\boldsymbol{P}_n$ as the class of probabilities π on $\mathcal{X}$ that assume only values of the form r/n for integral r.

THEOREM 12.1. *There is a* 1-to-1 *correspondence* $\tilde{\pi} \to \pi$ *carrying* $\tilde{\boldsymbol{P}}_n$ *onto* $\boldsymbol{P}_n$ *such that*

$$\pi E = \tilde{\pi} C(E) \tag{12.1}$$

for all $E \in \mathcal{X}$. *This correspondence is uniquely determined by condition* (12.1). *Under this correspondence, the* σ*-algebra generated by sets of the form* $\{\tilde{\pi} \mid \tilde{\pi} A \leqq \alpha\}$ *in* $\tilde{\boldsymbol{P}}_n$ *is carried* σ*-isomorphically onto the* σ*-algebra generated by sets of the form* $\{\pi \mid \pi E \leqq \epsilon\}$ *in* $\boldsymbol{P}_n$. *There are algorithms for computing* $\tilde{\pi}$ *from* π *which are given in detail in the proof.*

Proof. The set-function defined for all $E \in \mathcal{X}$ as $\tilde{\pi} C(E)$ is obviously a probability on $\mathcal{X}$. We shall show that this probability assumes only the values $0, 1/n, 2/n, \cdots, 1$. Let A_r be the set of points of $\tilde{X}_n$ for which exactly r coordinates lie in E. It is clear that A_r is measurable and invariant under G. Since $\tilde{\pi}$ is invariant, Theorem 11.4 implies that $\tilde{\pi} A_r = 0$ or 1. The sets $A_0, A_1, \cdots, A_n$ form a partition of $\tilde{X}_n$, and therefore there is a unique r $(0 \leqq r \leqq n)$ such that $\tilde{\pi} A_r = 1$. The set A_r is partitioned in a natural way into $C_{n,r}$ subsets, each characterized by the set of r co-ordinates that lie in E. Each of these subsets of A_r can be transformed into all of the others by a transformation of G, and hence all have equal $\tilde{\pi}$ measure, which must be $1/C_{n,r}$. If $r = 0$, then $\tilde{\pi} C(E) = 0$. If $r > 0$, then the set $A_r \cap C(E)$ is the union of $C_{n-1,r-1}$ of these subsets. Therefore,

$$\tilde{\pi} C(E) = \tilde{\pi}(A_r \cap C(E)) = C_{n-1,r-1}/C_{n,r} = r/n. \tag{12.2}$$

It follows that (12.1) defines a mapping of $\tilde{\boldsymbol{P}}_n$ into $\boldsymbol{P}_n$.

To show that the mapping defined by (12.1) carries $\tilde{\boldsymbol{P}}_n$ onto $\boldsymbol{P}_n$, we take advantage of the fact that every $\pi \in \boldsymbol{P}_n$ can be represented in the form

$$\pi = \frac{1}{n}\sum_{i=1}^{n} \pi_i, \tag{12.3}$$

where the π_i are 0-1 probabilities on $\mathfrak{X}$ (some π_i may be identical with some π_j for $i \neq j$). We leave to the reader the simple proof that the representation (12.3) is always possible and that the π_i are unique except for their order. We now write $\theta = \pi_1 \times \pi_2 \times \cdots \times \pi_n$, the product measure of $\pi_1, \cdots, \pi_n$, in that order, on $\tilde{\mathfrak{X}}_n$. Since every π_i is a 0-1 probability, it follows that θ is also a 0-1 probability. Defining σ_θ as in Theorem 11.5, we see from Theorem 11.6 that $\sigma_\theta \in \tilde{\boldsymbol{P}}_n$. An easy calculation shows now that

$$\sigma_\theta C(E) = \frac{1}{n}\sum_{i=1}^{n} \pi_i E = \pi E, \tag{12.4}$$

so that our mapping of $\tilde{\boldsymbol{P}}_n$ into $\boldsymbol{P}_n$ is in fact onto.

To show that the mapping described in (12.1) is 1-to-1, it suffices to show that $\tilde{\pi}$ is determined for all cylinders by the values of $\tilde{\pi} C(E)$ for all $E \in \mathfrak{X}$. We need consider only cylinders $C(E_1, \cdots, E_n)$ where, for all $i, j = 1, 2, \cdots, n$, E_i and E_j are identical or disjoint (since every cylinder is obviously the union of a finite number of cylinders each having this property). Now suppose that $F_1, \cdots, F_t$ are the distinct sets occurring among the sets $E_1, \cdots, E_n$ and that F_u appears r_u times ($u = 1, \cdots, t$). The cylinder $C(E_1, \cdots, E_n)$ then has exactly $n!/r_1! \cdots r_t!$ distinct transforms under the group G, and all of these are obviously disjoint.

Now let $A_{i,s}$ be the set of all $a \in \tilde{X}_n$ exactly s co-ordinates of which lie in the set F_i ($i = 1, \cdots, t$, $s = 0, 1, 2, \cdots, n$). As in the considerations leading to (12.2), we see that there exists exactly 1 s_i such that $\tilde{\pi}(A_{i,s_i}) = 1$ ($i = 1, \cdots, t$), and that accordingly

$$\tilde{\pi}(A_{1,s_1} \cap \cdots \cap A_{t,s_t}) = 1. \tag{12.5}$$

Also

$$\tilde{\pi} C(F_i) = s_i/n \qquad (i = 1, \cdots, t). \tag{12.6}$$

Now, as already noted, $C(E_1, \cdots, E_n)$ is one of $n!/r_1! \cdots r_t!$ elements of a partition of $A_{1,r_1} \cap \cdots \cap A_{t,r_t}$, all of these sets being congruent under transformations in G. Hence we have

$$\tilde{\pi} C(E_1, \cdots, E_n) = \begin{cases} r_1! \cdots r_t!/n! & \text{if } r_i = s_i \ (i = 1, \cdots, t), \\ 0 \text{ otherwise.} \end{cases} \tag{12.7}$$

Relations (12.6) and (12.7) show that $\tilde{\pi}$ is determined for all cylinders $C(E_1, \cdots, E_n)$ with disjoint or identical E_i by its values on cylinders $C(E)$, and thus the mapping described in (12.1) is 1-to-1 onto. This completes the proof of the first statement of the present theorem. The proof of the other

2 statements are now simple and are left to the reader.

From Theorems 12.1 and 11.7, we now draw the following conclusion.

THEOREM 12.3. *For every $\sigma \in \mathbf{S}_n$, there exists a probability μ on the σ-algebra of subsets of $\boldsymbol{P}_n$ generated by sets of the form*

$$\{\pi \mid \pi E \leqq \epsilon\} \qquad (E \in \mathfrak{X}, \epsilon \text{ real}) \tag{12.8}$$

such that

$$\sigma A = \int_{\boldsymbol{P}_n} (\tilde{}A)\, d\mu(\pi) \tag{12.9}$$

for all $A \in \tilde{\mathfrak{X}}_n$. The probability μ is uniquely determined by (12.9), *and can be computed by the algorithm*

$$\begin{aligned} &\mu\{\pi \mid \pi E_i \leqq \epsilon_i,\ i = 1, \cdots, t\} = \sigma\{a \mid a \in \tilde{X}_n, \\ &\quad \text{at most } n\epsilon_i \text{ co-ordinates of } a \text{ are in } E_i,\ i = 1, \cdots, t\}. \end{aligned} \tag{12.10}$$

Here $E_1, \cdots, E_t$ are disjoint elements of $\mathfrak{X}$ and $\epsilon_1, \cdots, \epsilon_t$ are real numbers.

Since product probabilities are exactly the extreme points of $\tilde{\mathbf{S}}$ (infinite product), but not of $\tilde{\mathbf{S}}_n$ in general, it is interesting to see what form (12.10) takes on when σ is an n-fold product of a probability π_0 on X. In this case, writing $F = X \cap (\bigcap_{i=1}^{t} E_i)'$, we have

$$\begin{aligned} &\mu\{\pi \mid \pi E_i \leqq \epsilon_i,\ i = 1, \cdots, t\} \\ &= \sum_{0 \leqq r_i \leqq \epsilon_i n,\ \Sigma r_i \leqq n} \frac{n!}{r_1! \cdots r_t! u!} (\pi_0 E_1)^{r_1} \cdots (\pi_0 E_t)^{r_t} (\pi_0 F)^u \end{aligned} \tag{12.11}$$

where $u = n - \sum r_i$. For large n, according to well-known properties of the multinomial distribution [6, p. 318], the right side of (12.11) is nearly 0 if some $\epsilon_i < \pi_0 E_i$ and is nearly 1 if all $\epsilon_i > \pi_0 E_i$. Thus, as might be expected, μ tends to be concentrated on probabilities π that closely resemble π_0.

The analogy between (2.3) and (12.9) will be made a little more complete by showing that the elements of $\tilde{\boldsymbol{P}}_n$ tend, in a sense, to resemble product measures for large n. Consider a $\tilde{\pi} \in \tilde{\boldsymbol{P}}_n$ corresponding to a $\pi \in \boldsymbol{P}_n$, which is made up of 0-1 probabilities π_i in the sense of (12.3). To see that $\tilde{\pi}$ does tend to resemble a product measure, we estimate $\tilde{\pi} C(E_1, \cdots, E_r)$, as follows.

$$\tilde{\pi} C(E_1, \cdots, E_r) = \frac{(n-r)!}{n!} \Sigma', \tag{12.12}$$

where

$$\Sigma' = \Sigma' \pi_{i_1} E_1 \cdot \pi_{i_2} E_2 \cdots \pi_{i_r} E_r, \tag{12.13}$$

the summation being restricted to terms such that $i_1, \cdots, i_r$ are all different. If Σ denotes the corresponding unrestricted sum, which is, of course, $n^r \pi E_1 \cdots \pi E_n$, then we have

$$
\begin{aligned}
&\left| n^{-r}\Sigma - \frac{(n-r)!}{n!}\Sigma' \right| \\
&\qquad \leqq n^{-r}\left|\Sigma - \Sigma'\right| + \left| n^{-r} - \frac{(n-r)!}{n!}\right|\Sigma' \\
&\qquad \leqq n^{-r}\left(n^r - \frac{n!}{(n-r)!}\right) + \left(\frac{(n-r)!}{n!} - n^{-r}\right)\frac{n!}{(n-r)!} \\
&\qquad = 2\left(1 - \frac{n!}{n^r(n-r)!}\right) \\
&\qquad \leqq 2\left\{1 - \left(1 - \frac{r-1}{n}\right)\right\} = O\left(\frac{1}{n}\right),
\end{aligned}
\tag{12.14}
$$

so $\tilde{\pi}$ applied to any set determined by a few co-ordinates only is for large n almost the same as the iterated product of π applied to the same set.

References

1. P. S. Alexandroff and H. Hopf, *Topologie* I. Die Grundlehren der mathematischen Wissenschaften, vol. XLV, Berlin, Springer, 1935.

2. Paul Alexandroff and Paul Urysohn, *Mémoire sur les espaces topologiques compacts*, Verhandelingen der Koninklijke Akademie van Wetenschappen te Amsterdam, Afdeeling Natuurkunde (Eerste Sectie) vol. XIV, no. 1, Amsterdam, 1929.

3. T. Bonnesen and W. Fenchel, *Theorie der konvexen Körper*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Berlin, 1934. (Reprint, New York, Chelsea, 1948.)

4. N. Bourbaki, *Éléments de mathématique.* First part, Book V, *Espaces vectoriels topologiques*, Chapters I–II, Actualités Scientifiques et Industrielles, no. 1189, Paris, Hermann, 1953.

5. ———, *Éléments de mathématique.* First part, Book VI, *Intégration*, Chapters I–IV, Actualités Scientifiques et Industrielles, no. 1175, Paris, Hermann, 1952.

6. Harald Cramér, *Mathematical methods of statistics*, Princeton University Press, 1951.

7. J. L. Doob, *Application of the theory of martingales. Le calcul des probabilités et ses applications*, Coll. Intern. du Centre de la Recherche Scientifique, no. 13, pp. 23–27, Paris, Centre Nat. Rech. Sci., 1949.

8. ———, *Stochastic processes*, New York, Wiley, 1953.

9. E. B. Dynkin, *Klassy ekvivalentnyh Slučaĭnyh veličin*, Uspehi Matematičeskih Nauk 8 vol. 54 (1953) pp. 125–134.

10. Bruno de Finetti, *Funzione caratteristica di un fenomeno aleatorio*, Atti della R. Accademia Nazionale dei Lincei, Ser. 6, Memorie, Classe di Scienze Fisiche, Matematiche e Naturali vol. 4 (1931) pp. 251–299.

11. ———, *La prévision: ses lois logiques, ses sources subjectives*, Annales de l'Institut Henri Poincaré vol. 7 (1937) pp. 1–68.

12. Maurice Fréchet, *Les probabilités associées à un système d'événements compatibles et dépendants.* II, Actualités Scientifiques et Industrielles, no. 942, Paris, Hermann, 1943.

13. Jules Haag, *Sur un problème général de probabilités et ses diverses applications*, Proceedings of the International Congress of Mathematicians, Toronto, 1924, Toronto, 1928, pp. 659–674.

14. Paul Halmos, *Measure theory*, New York, Van Nostrand, 1950.

15. Felix Hausdorff, *Grundzüge der Mengenlehre*, Leipzig, Viet, 1914. (Reprinted, New York, Chelsea, 1949.)

16. Edwin Hewitt, *Linear functionals on spaces of continuous functions*, Fund. Math. vol. 37 (1950) pp. 161–189.

17. Edwin Hewitt and L. J. Savage, Bull. Amer. Math. Soc. Abstract 59-4-523.

18. A. Ya. Hinčin, *O klassah ekvivalentnyh sobytiĭ*, Doklady Akad. Nauk SSSR vol. 85 (1952) pp. 713–714.

19. ———, *Sur les classes d'événements équivalents*, Mat. Sbornik vol. 39 (1932) pp. 40–43.

20. M. G. Kreĭn and D. Mil'man, *On extreme points of regular convex sets*, Studia Mathematica vol. 9 (1940) pp. 133–139.

21. Lynn H. Loomis, *On the representation of σ-complete Boolean algebras*, Bull. Amer. Math. Soc. vol. 53 (1947) pp. 757–760.

22. John C. Oxtoby, *Ergodic sets*, Bull. Amer. Math. Soc. vol. 58 (1952) pp. 116–136.

23. B. J. Pettis, *On integration in vector spaces*, Trans. Amer. Math. Soc. vol. 44 (1938) pp. 277–304.

24. L. J. Savage, *The foundations of statistics*, New York, Wiley, 1954.

25. M. H. Stone, *Applications of the theory of Boolean rings to general topology*, Trans. Amer. Math. Soc. vol. 41 (1937) pp. 375–481.

26. ———, *The theory of representations for Boolean algebras*, Trans. Amer. Math. Soc. vol. 40 (1936) pp. 37–111.

27. Minoru Tomita, *On the regularly convex hull of a set in a conjugate Banach space*, Math. J. Okayama Univ. vol. 3 (1954) pp. 143–145.

28. Kôsaku Yosida and Edwin Hewitt, *Finitely additive measures*, Trans. Amer. Math. Soc. vol. 72 (1952) pp. 46–66.

THE UNIVERSITY OF WASHINGTON,
SEATTLE, WASH.
THE UNIVERSITY OF CHICAGO,
CHICAGO, ILL.